PROOF THAT CERTAIN IDEALS IN A CYCLOTOMIC REALM ARE PRINCIPAL IDEALS*

BY

HOWARD H. MITCHELL

1. Introduction

A theorem which serves to establish a connection between cyclotomy and the representation of numbers by quadratic forms is as follows:

Let λ represent a prime of the form $4n+3$, and q a prime that is congruent to 1, mod λ. Further, let Σa, Σb represent the sums of those integers that are respectively quadratic residues and non-residues of λ and that lie between 0 and λ. Then it is always possible to find integers x, y such that

$$x^2 + \lambda y^2 = 4q^{(\Sigma b - \Sigma a)/\lambda}.$$

This theorem, which seems to have been discovered independently by Jacobi and Cauchy,† was extended by the latter to cases where λ is a composite integer of certain types.

An essentially equivalent theorem is that if $\mathfrak{q}$ be a prime ideal factor of q in the quadratic realm $k(\sqrt{-\lambda})$, then

$$\mathfrak{q}^{(\Sigma b - \Sigma a)/\lambda}$$

is a principal ideal.‡

This result was later extended by Eisenstein§ and Stickelberger‖ to cases where $q \not\equiv 1$, mod λ. When λ is prime, the theorem holds provided q is a quadratic residue of λ.

These results follow also from Dirichlet's determination of the number of classes of quadratic forms of determinant $-\lambda$.¶

* Presented to the Society, December 28, 1916.

† Jacobi, *Werke*, vol. 6, pp. 233, 240, 254, 275; Cauchy, *Oeuvres*, 1 Serie, vol. 5, pp. 52, 64, 85, 95; H. J. S. Smith, *Report on the Theory of Numbers, Collected Mathematical Works*, vol. 1, pp. 273–283; Bachmann, *Die Lehre von der Kreistheilung* (1872), p. 290.

‡ Hilbert, *Die Theorie der algebraischen Zahlkörper*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 4 (1894–95), p. 386; *Encyclopädie der Mathematischen Wissenschaften*, Bd. I, p. 703; *Encyclopédie des Sciences Mathématiques*, Tome I, volume 3, p. 448.

§ Journal für Mathematik, vol. 37 (1848), pp. 97–126.

‖ Mathematische Annalen, vol. 37 (1890), p. 360.

¶ Dirichlet-Dedekind, *Zahlentheorie*, 4th edition, § 104. For a determination of the closely related class number for ideals see Hilbert, l. c., p. 320.

The author obtains a theorem in this paper which may be regarded as a further extension of these results (for the case of λ prime) to realms of higher degree. Let λ denote any prime such that $\lambda - 1$ is not a power of 2, and let $2e$ denote a divisor of $\lambda - 1$ such that $(\lambda - 1)/(2e)$ is an odd integer other than unity. Let q be a prime such that

$$q^{(\lambda-1)/2e} \equiv 1, \qquad \mod \lambda.$$

Then in the realm of degree $2e$ determined by the $2e$ periods formed from λth roots of unity it is known that q is the product of $2e$ conjugate prime ideals, which form e conjugate imaginary pairs. Further, let h denote the "first factor" of the class number of the realm.* Then if one ideal be selected from each of the e conjugate imaginary pairs, and the product of these be raised to the power h, it will be shown that the result is a principal ideal.

Expressed in another way, this theorem would state that in whatever way q^h be expressed as the product of two conjugate imaginary factors that are relatively prime, the two factors are principal ideals. It then follows from a property of units that q^h may be expressed in exactly 2^{e-1} ways as the product of two actual conjugate imaginary factors that are relatively prime, provided two factors that differ merely in sign be regarded as the same.

Another result is that if two conjugate imaginary prime ideal factors of q be each raised to the power h, the ideals thus obtained belong to the same ideal class.

2. The modified Kummer function $\psi(\alpha)$

We consider the cyclotomic function of Jacobi and Cauchy as generalized by Kummer,†

$$\psi_{-a,-b}(\alpha) = \sum_{\tau} \alpha^{-b \operatorname{ind} \tau + (a+b) \operatorname{ind} (\tau+1)},$$

where α is a primitive λth root of unity, a, b integers such that $ab(a+b) \not\equiv 0$, mod λ, and the summation is to be taken over all the marks τ (except 0 and -1) of a Galois field of order q^t, where ind τ means the index with respect to a primitive root in the Galois field, and t is the smallest exponent such that $q^t \equiv 1$, mod λ.

This function has the properties:

$$\psi(\alpha^q) = \psi(\alpha),$$

$$\psi(\alpha)\psi(\alpha^{-1}) = q^t.$$

* Kummer, Journal für Mathematik, vol. 40 (1850), p. 112. Fuchs has given an extension to cases where λ is composite, ibid., vol. 65 (1866), pp. 102, 106.

† For the case where $q \equiv 1$, mod λ, see for example H. Weber, *Lehrbuch der Algebra* (1898), vol. 1, § 177, 178; vol. 2, § 203; for the general q, Kummer, Journal für Mathematik, vol. 44 (1851), pp. 106–121; Stickelberger, l. c., p. 335; the author, these Transactions, vol. 17 (1916), pp. 165–177.

We shall suppose that λ is prime and that t is odd. In view of the first of these properties the function belongs to the subrealm of the realm $k(\alpha)$ that is determined by the $(\lambda - 1)/t$ periods. In this realm q is the product of $(\lambda - 1)/t$ prime ideal factors, which are also prime in the realm $k(\alpha)$.

Aside from an error that has been corrected by the author,* Kummer determined the prime ideal factors of this function. Let $\mathfrak{q}_0 = \mathfrak{q}_0(\alpha)$ denote any one of the prime ideal factors of q, and let

$$\mathfrak{q}_i = \mathfrak{q}_0(\alpha^{\gamma^i}),$$

where γ denotes a primitive root of λ, and the index i is to be reduced, mod $(\lambda - 1)/t$. Then if we write

$$[\psi(\alpha)] = \prod_i \mathfrak{q}_i^{m_i},$$

and if the primitive root in the Galois field of order q^t be properly selected, the exponents m_i are determined by the equations,

$$m_i = S_{-i+\operatorname{ind} a} + S_{-i+\operatorname{ind} b} - S_{-i+\operatorname{ind}(a+b)},\dagger$$

where λS_j denotes the sum of all the positive integers between 0 and λ whose indices with respect to the primitive root γ are congruent to j, mod $(\lambda - 1)/t$.

We now denote by $2e$ any divisor of $\lambda - 1$ such that $(\lambda - 1)/2e$ is an odd integer other than unity, and assume that

$$q^{(\lambda-1)/2e} \equiv 1, \qquad \operatorname{mod} \lambda.$$

In case $t = (\lambda - 1)/2e$, the function $\psi(\alpha)$ belongs to the realm determined by the $2e$ periods and the ideals $\mathfrak{q}_i$ are the prime ideal factors of q in that realm. Hence

$$\prod_i \mathfrak{q}_i^{m_i} \qquad (i = 0, 1, \cdots, 2e - 1)$$

is a principal ideal.

We shall show that a similar statement holds in case t is a divisor of $(\lambda - 1)/2e$, say $t = (\lambda - 1)/2e'$, where e' is a multiple of e. In this case we consider the product of $\psi(\alpha)$ by those of its conjugates obtained by replacing α by

$$\alpha^{\gamma^{2e}}, \ \alpha^{\gamma^{4e}}, \ \cdots, \ \alpha^{\gamma^{2e'-2e}}.$$

This product is contained in the realm determined by the $2e$ periods. The exponents of the e'/e ideals,

$$\mathfrak{q}_i, \ \mathfrak{q}_{i+2e}, \ \cdots, \ \mathfrak{q}_{i+2e'-2e},$$

* L. c., p. 171.

† This expression may be obtained either from that given by the author (l. c., p. 173) by replacing i by γ^{-i}, or else from that given by Kummer (l. c., p. 120) in case the function is replaced by the proper one of its conjugates.

which are conjugate under these substitutions are each equal to

$$m_i + m_{i+2e} + \cdots + m_{i+2e'-2e}.$$

This sum may itself be represented in the form represented by m_i above provided λS_j is now considered to represent the sum of the integers between 0 and λ, whose indices are congruent to j, mod $2e$. Moreover the product of these e'/e conjugate ideals is a prime ideal in the realm determined by the $2e$ periods. If we represent it by $\mathfrak{q}_i$, where the index is now to be reduced, mod $2e$, and where $\mathfrak{q}_i = \mathfrak{q}_0(\alpha^{\gamma^i})$, we conclude that

$$\prod_i \mathfrak{q}_i^{m_i} \qquad (i = 0, 1, 2, \cdots, 2e-1)$$

is a principal ideal.

We now suppose that a, b are so chosen that

$$\text{ind } a \equiv 0, \qquad \text{ind } (a+b) \equiv \text{ind } b, \qquad \text{mod } 2e.$$

Since we are assuming that $(\lambda - 1)/2e$ is greater than 1, we may select them in such a way that the second condition is satisfied. Then by multiplying each of them by a properly selected residue, mod λ, the second condition will continue to hold, and the first condition can be satisfied. With these conditions imposed on a, b the exponents of the ideals assume the special values, $m_i = S_{-i}$.

3. Proof of the theorem

The functions conjugate to that considered above under the substitutions in which α is replaced by

$$\alpha^{\gamma}, \ \alpha^{\gamma^2}, \ \cdots, \ \alpha^{\gamma^{e-1}}$$

have the following expressions in terms of the prime ideal factors

$$\prod_i \mathfrak{q}_i^{m_{i-1}}, \ \prod_i \mathfrak{q}_i^{m_{i-2}}, \ \cdots, \ \prod_i \mathfrak{q}_i^{m_{i-e+1}}.$$

If now we raise the original function to the power p_0, and its conjugates to the powers $p_1, p_2, \cdots, p_{e-1}$ respectively, and then form the product, the exponent of $\mathfrak{q}_i$ in the product will be equal to

$$p_0 m_i + p_1 m_{i-1} + \cdots + p_{e-1} m_{i-e+1}.$$

If we equate these exponents for $i = 0, 1, \cdots, e-1$, we obtain $e-1$ equations which we write as follows:

$$p_0(m_i - m_0) + p_1(m_{i-1} - m_{-1}) + \cdots + p_{e-1}(m_{i-e+1} - m_{-e+1}) = 0 \qquad (i = 1, 2, \cdots, e-1).$$

In view of our assumption that $(\lambda - 1)/2e$ is odd, ind $(-1) \equiv e$, mod $2e$, and hence if a number r appears in the sum λS_j, $\lambda - r$ will appear in the

sum λS_{j+e}. Consequently

$$m_i + m_{i+e} = (\lambda - 1)/2e.$$

We set

$$m_i - m_{i+e} = d_i,$$

where d_i is an odd integer, and write the above equations in the form

$$p_0(d_i - d_0) + p_1(d_{i-1} - d_{-1}) + \cdots + p_{e-1}(d_{i-e+1} - d_{-e+1}) = 0.$$

If the exponents of the ideals $\mathfrak{q}_i$ be equated for $i = e, e+1, \cdots, 2e-1$, the same set of equations are obtained. Hence if rational integers $p_0, p_1, \cdots, p_{e-1}$ can be found to satisfy these equations, the product formed as described will assume the form

$$(\mathfrak{q}_0 \mathfrak{q}_1 \cdots \mathfrak{q}_{e-1})^k (\mathfrak{q}_e \mathfrak{q}_{e+1} \cdots \mathfrak{q}_{2e-1})^{k-h},$$

where k and h are integers, and hence

$$(\mathfrak{q}_0 \mathfrak{q}_1 \cdots \mathfrak{q}_{e-1})^h$$

will be a principal ideal.

These equations will all be satisfied if we take for the integers p_j the values,

$$p_j = D_j/2^{e-1},$$

where D_j denotes the minor of the element d_{-j} in the first row of the determinant

$$D = \begin{vmatrix} d_0 & d_{-1} & \cdots & d_{-e+1} \\ d_1 - d_0 & d_0 - d_{-1} & \cdots & d_{-e+2} - d_{-e+1} \\ d_2 - d_0 & d_1 - d_{-1} & \cdots & d_{-e+3} - d_{-e+1} \\ \vdots & \vdots & & \vdots \\ d_{e-1} - d_0 & d_{e-2} - d_{-1} & \cdots & d_0 - d_{-e+1} \end{vmatrix}.$$

The exponent h is equal to

$$p_0 d_0 + p_1 d_{-1} + \cdots + p_{e-1} d_{-e+1},$$

and if the p's have the values given above, this is equal to

$$h = D/2^{e-1}.$$

In view of the relations, $d_{-i} = -d_{e-i}$, the determinant D, if the first row be added to each of the others, belongs to the class known as "skew circulants,"* and may thus be expressed in the form

$$D = \prod_i (d_0 + d_{-1}\beta^i + d_{-2}\beta^{2i} + \cdots + d_{-e+1}\beta^{(e-1)i}),$$

where β is a primitive root of the equation $x^e = -1$, and the product is to be taken over the values $i = 1, 3, \cdots, 2e-1$.

* See, for example, Muir, *Theory of Determinants* (1882), pp. 187–191.

Since

$$d_{-j} = m_{-j} - m_{-j-e} = S_j - S_{j+e} = S_j + S_{j+e}\,\beta^{ei},$$

provided i is odd, we may write

$$h = \frac{1}{2^{e-1}} \prod_i (S_0 + S_1\,\beta^i + S_2\,\beta^{2i} + \cdots + S_{2e-1}\,\beta^{(2e-1)i}).$$

This expression is equal to the "first factor" of the class number of the realm determined by the $2e$ periods.*

This result may be made more general by observing that if instead of equating the exponents of the ideals $\mathfrak{q}_0, \mathfrak{q}_1, \cdots, \mathfrak{q}_{e-1}$, we had replaced part or all of them by their conjugate imaginary ideals, and followed the same procedure, the same value of h would have been obtained except possibly for sign. The only difference in the determinant D, when simplified by the addition of the first row to each of the others, would have been that the signs of all the elements in some of the rows would have been changed.

We therefore conclude that in whatever way q^h be represented as the product of two conjugate imaginary ideal factors that are relatively prime, those factors are principal ideals.

4. Some consequences and an example

There are two consequences of this theorem that may be mentioned. If in the expression

$$(\mathfrak{q}_0\,\mathfrak{q}_1 \cdots \mathfrak{q}_{e-1})^h$$

we replace one of the prime ideals $\mathfrak{q}_i$ by its conjugate imaginary $\mathfrak{q}_{i+e}$, and then form the quotient of the two expressions, we conclude that $\mathfrak{q}_i^h/\mathfrak{q}_{i+e}^h$ may be expressed as the quotient of two principal ideals and hence that $\mathfrak{q}_i^h$ and $\mathfrak{q}_{i+e}^h$ belong to the same ideal class.

Another consequence is that q^h may be represented in exactly 2^{e-1} ways as the product of two actual conjugate imaginary factors that are relatively prime, provided we regard as the same two factors that differ merely in sign. For if

$$q^h = F(\alpha)\,F(\alpha^{-1}),$$

and if $E(\alpha)$ denotes a unit such that

$$q^h = F(\alpha)\,E(\alpha)\,F(\alpha^{-1})\,E(\alpha^{-1}),$$

then $E(\alpha)\,E(\alpha^{-1}) = 1$. By a theorem due to Kummer† it follows from this that $E(\alpha)$ is a root of unity. But the only roots of unity in the realm determined by the $2e$ periods are ± 1.

* L. c.

† Extensions have been given by Kronecker and Minkowski; see Hilbert, l. c., p. 221.

As an example consider the case where $\lambda = 31$, $2e = 6$, $q = 2$. We select the primitive root γ of 31 to be 3, and represent by η_i the sum of the five 31st roots α^j such that $\text{ind}_3 j \equiv i$, mod 6. We find that $a = 1$, $b = 1$ satisfy the conditions,

$$\text{ind } a \equiv 0, \quad \text{mod } 6; \qquad \text{ind } b \equiv \text{ind } (a + b), \quad \text{mod } 6.$$

If then we select the primitive root in the Galois field of order 2^5 to be a root of the irreducible congruence

$$x^5 + x^2 + 1 \equiv 0, \quad \text{mod } 2,$$

we find that

$$\psi_{-1,-1}(\alpha) = \sum_{\tau} \alpha^{-\text{ind } \tau + 2 \text{ ind } (\tau+1)} = 2(\eta_0 + \eta_1 + \eta_3),$$

where the summation is taken over all the marks τ of the field except 0 and 1.

If $\mathfrak{q}_0$ is the properly selected prime ideal factor of 2, the exponents of the ideals $\mathfrak{q}_i$ are $m_i = S_{-i}$, and these are found to be respectively

$$1, 3, 3, 4, 2, 2$$

for $i \equiv 0, 1, 2, 3, 4, 5$, mod 6. The exponents of the ideal factors in the two conjugate functions obtained by replacing α by α^3 and α^9 are

$$2, 1, 3, 3, 4, 2,$$
$$2, 2, 1, 3, 3, 4.$$

Hence the exponents of the ideal factors in the product

$$\psi(\alpha)\psi^2(\alpha^3)\psi^4(\alpha^9)$$

are

$$13, 13, 13, 22, 22, 22.$$

We find that this product is equal to

$$2^{13}(-6 - \eta_0 + 10\eta_1 + 4\eta_2 + 5\eta_4 + \eta_5),$$

and we thus obtain an expression for 2^9 as the product of two conjugate imaginary factors, i. e.,

$$2^9 = (-6 - \eta_0 + 10\eta_1 + 4\eta_2 + 5\eta_4 + \eta_5) \times (-6 - \eta_3 + 10\eta_4 + 4\eta_5 + 5\eta_1 + \eta_2).$$

That these factors are relatively prime may be verified by noticing that their difference may be expressed in the form

$$1 + 6\eta_1 + 4\eta_2 + 2\eta_3 - 4\eta_4 - 2\eta_5,$$

which is obviously prime to 2. We readily find that 9 is the first factor of the class number of the realm under consideration.

Three other such factorizations of 2^9 are possible, two of which may be obtained from that given above by permuting the η's. In the other the factors would be expressible in terms of $\sqrt{-31}$, a case that has been fully treated by Stickelberger,* since, if the included quadratic realm is imaginary, its class number divides the first factor of the class number of the realm determined by the $2e$ periods. The class number of $k(\sqrt{-31})$ is 3, and we have obviously

$$2^3 = \frac{1+\sqrt{-31}}{2} \cdot \frac{1-\sqrt{-31}}{2}.$$

UNIVERSITY OF PENNSYLVANIA,
PHILADELPHIA, PA.

* L. c.

THE ORDER OF PRIMITIVE GROUPS* (III)

BY

W. A. MANNING

1. The theorem which I now propose to prove is

THEOREM XIII. *Let q be any positive integer greater than 5, and let p be any prime greater than $2q-2$; then the degree of any primitive group G that contains a substitution of order p and degree qp but none of order p and of degree less than qp does not exceed $qp+4q-4$; if G is not triply (doubly) transitive its degree does not exceed $qp+4q-6$ ($qp+4q-7$); the order of G is not divisible by p^2.*

2. In the two former papers in these Transactions† under this same title, and to which references are indicated by the Roman numerals I and II in parenthesis, a proof and a slight extension were given (for q greater than unity) of the theorem which Jordan stated in the *Memoir on Primitive Groups* in the first volume of the Bulletin of the Mathematical Society of France, page 175:

Let q be a positive integer less than 6; p any prime greater than q; the degree of a primitive group G that contains a substitution of order p on q cycles (without including the alternating group) can not exceed $qp+q+1$.

To this one may add that when p is greater than $q+1$, the order of G is not divisible by p^2.

3. The memoir at the end of which this theorem is given is devoted to the larger question of the corresponding limit for the degree of G when q is not confined to numbers less than 6. Jordan's general result may be stated thus:

Let q be any positive integer, p any prime greater than $2q\log_2 q+q+1$; the degree of a primitive group G that contains a substitution of order p on q cycles (without including the alternating group) can not exceed $qp+2q\log_2 2q$.

4. This is supplemented in certain directions by the theorem:‡

If a primitive group of class greater than 3 contains a substitution of prime order p and of degree qp (q less than $2p+3$), it includes a transitive subgroup of degree not greater than the larger of the two numbers $qp+q^2-q$, $2q^2-p^2$. In particular, if q is less than $p+2$, the degree of G, when it is simply transitive, is not greater than $qp+q^2-q$.

* Presented to the Society, April 7, 1917.

† These Transactions, vol. 10 (1909), pp. 247–258; vol. 16 (1915), pp. 139–147. Theorem X asserts that the upper limit of the degree is $qp+q$ when p is greater than $q+1$, and q is greater than 1 and less than 5.

‡ These Transactions, vol. 12 (1911), pp. 375–386.

5. Simply transitive primitive groups are known which contain a substitution of prime order p and of degree qp $(q < p)$ and whose degree is $qp + q(q-2)/8$. For example, if the alternating group of degree n is represented as a transitive group on the $n(n-1)/2$ binary products ab, ac, $\cdots$, the number of these products left fixed by a circular substitution of prime order p is exactly $(n-p)(n-p-1)/2$. Then the given substitution of order p in the new representation of the alternating group is of degree $[n-(p+1)/2]p$; the condition that the number of cycles in this substitution is less than p, $n-(p+1)/2 < p$, may be written $n < 3p/2$. The number of letters left fixed by the given substitution of order p has therefore its maximum value when $n = (3p-1)/2$, viz., $(p-1)(p-3)/8$, when $n-(p+1)/2 = p-1$. Now

$$n(n-1)/2 = (p-1)p + (p-1)(p-3)/8 = qp + q(q-2)/8.$$

If

$$n = 3p/2 - (2k-1)/2, \qquad n-(p+1)/2 = p-k,$$

$$(n-p)(n-p-1)/2 = (p-2k+1)(p-2k-1)/8,$$

and we have

$$n(n-1)/2 = (p-k)p + (p-2k+1)(p-2k-1)/8,$$

$k = 1, 2, \cdots, (p-3)/2$, as the degree of a simply transitive primitive group which contains a substitution of order p and degree qp, $q < p$.

It is not improbable that the true limit $(q < p)$ is $qp + q(q-2)/8$ instead of $qp + q^2 - q$ but we know so few primitive groups that an induction from those known has not much value. At any rate we need not expect to extend the formula $qp + q + 1$ to all primitive groups in which q is less than p.

6. Another related theorem is that of Bochert:*

The class (> 3) *of a substitution group of degree* n *exceeds* $n/3 - 2\sqrt{n}/3$ *if it is doubly,* $n/3 - 1$ *if triply,* $n/2 - 2$ *if quadruply, transitive.*

7. This theorem, with that of Sylow,† is indispensable in the proof of Theorem XIII. Another theorem which has been of constant use in the two preceding numbers, and without which the present development would be well-nigh, probably quite, impossible, is

THEOREM XIV. *The largest subgroup of a transitive group* G *of degree* n, *in which a subgroup* H *that leaves fixed* m $(0 < m < n)$ *letters is invariant, has as many transitive constituents in these* m *letters as there are different conjugate sets in* G_1 (*a subgroup of* G *that leaves one of the* m *letters fixed*) *which, under the substitutions of* G, *enter into the complete set of conjugates to which* H *belongs. Moreover, the degree of each of these constituents is proportional to the number of subgroups in the several conjugate sets of* G_1 *in question.*

* Bochert, Mathematische Annalen, vol. 40 (1892), pp. 176–193; vol. 49 (1897), pp. 133–144.

† Sylow, Mathematische Annalen, vol. 5 (1873), pp. 584–594.

A clear comprehension of this theorem is so necessary to the reader that I venture to insert here a proof of it, considerably fuller than the brief indication given in the Bulletin of the American Mathematical Society.*

8. Let g be the order of G, and let those conjugates of H, which are found in G_1, lie in k different sets, in so far as they are permuted by the substitutions of G_1 only, with r conjugates in the set that includes H, r_1 in a second set, and so on. In G_1, H is invariant in a subgroup of order g/nr, while H_1, a subgroup in the second set, is invariant in a group of order g/nr_1. The largest subgroup I of G in which H is invariant is of order gm/ns, $s = r + r_1 + \cdots + r_{k-1}$. Now I does not connect transitively the $n - m$ letters displaced by H and the m letters it leaves fixed. Since the largest subgroup of G_1 in which H is invariant is of order g/nr, I has one transitive constituent of degree $(gm/ns)/(g/nr) = mr/s$ in letters left fixed by H. Let $a_1, a_2, \cdots$, $b_1, b_2, \cdots$ be the letters of G left fixed by H. The letter fixed by G_1 is a_1. Consider a substitution $S = a_1 b_1 \cdots$ of G. Since there are substitutions of G which transform H_1 into H we may assume that S transforms H_1 into H. Then $SHS^{-1} = H_1$, $(G_1 S) H (G_1 S)^{-1} = G_1 H_1 G_1^{-1} = H_1'$, that is, every substitution $G_1 S = a_1 b_1 \cdots$ of G transforms some subgroup H_1' (conjugate to H_1 under substitutions of G_1) into H: $(G_1 S)^{-1} H_1' (G_1 S) = H$. Then no matter by which of the substitutions $a_1 b_1 \cdots$ of G we transform G_1 we get a group G_2 that fixes b_1 and in which H is a member of that set of r_1 conjugate subgroups by which every substitution $G_1 S = a_1 b_1 \cdots$ replaces the set $H_1, \cdots$ of G_1. Then b_1 is one of the letters of a transitive constituent of degree mr_1/s in I. Since no substitution $a_1 b_1 \cdots$ transforms H into itself, the letter a_1 is not an element of this transitive constituent in the mr_1/s letters $b_1, b_2, \cdots$. If mr_1/s is not unity, there is a substitution $T_2 = b_1 b_2 \cdots$ in I, and the product $ST_2 = a_1 b_2 \cdots$ transforms H_1 into H, as do also the products $ST_3 = a_1 b_3 \cdots$, $\cdots$, where $T_3 = b_1 b_3 \cdots$, $\cdots$. Every substitution $G_1 ST_i$, $i = 2, 3, \cdots$, transforms some subgroup H_1' (conjugate of H_1 in G_1) into H. Suppose that a substitution $S_1 = a_1 c_1 \cdots$ transforms H_1 into H. Then $S^{-1} = b_1 a_1 \cdots$ transforms H into H_1, so that $S^{-1} S_1 = b_1 c_1 \cdots$ transforms H into itself, thereby showing that c_1 is one of the letters $b_1, b_2, \cdots$. If S_1 had transformed H_1' (some other member of the set H_1 of G_1) into H, a properly chosen substitution $G_1 S_1 = a_1 c_1 \cdots$ would have transformed H_1 into H, and that substitution could have been called S_1. Then no substitution $a_1 c_1 \cdots$ (c_1 not one of the letters $b_1, \cdots$) can transform any member of the set H_1 into H. Next, there must be a substitution in G to transform H_2 into H: call it $U = a_1 c_1 \cdots$. We know that c_1 does not belong to the same transitive constituent in I as a_1 or b_1. Then just as before I has a transitive constituent in the mr_2/s letters $c_1, c_2, \cdots$ associated with the conjugate set $H_2, \cdots$ of G_1.

* Vol. 13 (1906), p. 20.

Thus we find $k-1$ transitive constituents in $r_1 m/s$ letters $b_1, b_2, \cdots$, in $r_2 m/s$ letters $c_1, c_2, \cdots$, in $r_3 m/s$ letters $d_1, d_2, \cdots$, and so on, associated with the $k-1$ conjugate sets $H_1, \cdots, H_2, \cdots, H_3, \cdots, \cdots$ of G_1 in addition to the constituent of degree rm/s in the letters $a_1, a_2, \cdots$. Every substitution $a_1 a_2 \cdots, a_1 a_3 \cdots, \cdots$ transforms some subgroup H' (of the set $H, \cdots$ of G_1) into H.

We now pass to the consideration of Theorem XIII, which will be proved after the necessary preparation by a complete induction. The reader would do well to have freshly in mind the arguments of the two preceding numbers of this series on *the Order of Primitive Groups* and is recommended also to fortify himself by a perusal of the first twenty or thirty pages of Jordan's great *Memoir on Primitive Groups*. It will be seen that in setting up the subgroups H_{ij} in § 10 I have effected a combination of Jordan's method with that of my two earlier papers.

The subgroups H_{ij}

9. The group G is by hypothesis a primitive group in which there is a substitution of order p and degree qp but no substitution of the same order and of lower degree. From the beginning we assume that q is less than p, and ultimately we shall introduce the condition that p is greater than $2q-3$, but not before it seems unavoidable.

10. Let H_i be a subgroup of G that is generated by the similar substitutions $A_1, A_2, \cdots, A_i, A_1', A_2', \cdots, A_i', A_1'', A_2'', \cdots, A_i'', \cdots$, of order p and degree qp. It is to be understood as a part of the definition of H_i that all the substitutions of G of order p and degree qp and which are not in H_i displace one or more letters new to H_i.

If H_i is intransitive there exists in G a substitution A_{i+1}, similar to A_1, that unites two or more of the transitive sets of H_i (I, Theorem I). It is legitimate to assume that no other substitution similar to A_1 that joins two of the transitive sets of H_i displaces fewer new letters than does A_{i+1}. This being granted, A_{i+1} has at most one new letter in any cycle (I, Theorem IV). To $\{H_i, A_{i+1}\}$ are now to be adjoined all the other substitutions of G on the same letters that are similar to A_1. Call the resulting group H_{i+1}. It may be that there are in H_{i+1} and not in H_i certain substitutions of order p and degree qp which displace fewer letters new to H_i than A_{i+1} does. If so, no one of them connects sets of H_i. Let then $H_{ij}, j = 1, 2, \cdots (H_{i0} = H_i)$, be a subgroup of H_{i+1} that includes $H_i, H_{i1}, H_{i2}, \cdots, H_{i,j-1}$, and which is generated by $H_{i,j-1}$, a substitution B_{ij}, similar to A_1, which displaces a minimum number of letters new to $H_{i,j-1}$, hence at most one to any cycle, and all other substitutions of G of order p and degree qp on letters of $\{H_{i,j-1}, B_{ij}\}$ only. Let H_{ik_i} be the last of these groups before H_{i+1}. Let $x_1, x_2, \cdots,$

x_m ($m > 0$) be the letters of H_{i+1} that are not displaced by H_{ik_i}. Any substitution of H_{i+1} that replaces one of the m letters $x_1, x_2, \cdots, x_m$ by one of the same letters permutes these m letters only among themselves. This holds true of any two successive groups, as H_{ij} and $H_{i,\,j-1}$. Those of the letters $x_1, x_2, \cdots, x_m$ of H_{i+1} that belong to a transitive constituent of H_{i+1} form a system of imprimitivity of that constituent if their number exceeds unity. Since the number of letters in any system of imprimitivity of a group generated by substitutions of order p and degree less than p^2 is less than p, no letter x_1, say, can be associated in a system of imprimitivity with any letter of H_{ik_i}. It follows that a transitive constituent of H_{i+1} with just one new letter in it is primitive. This holds not only for H_{i+1} but for any group of the series, as H_{ij}.

11. For some value of i not greater than q, H_i is a transitive group. Let H_{r+1} be the first transitive group in the series. The group H_{r+2} is formed by the adjunction of a substitution A_{r+2} which displaces a minimum number of letters new to H_{r+1} and then all the substitutions of order p and degree qp in G on the letters of $\{H_{r+1}, A_{r+2}\}$. Of course A_{r+2} displaces at least one new letter if H_{r+1} is not the last group in the series $H_1, H_2, \cdots$. All the groups H_{r+1}, $H_{r+2}, \cdots$ are transitive. Finally a group H_{r+i} will be reached which is invariant in G.

12. When the degree of H_{r+1}, the first transitive group in the series, does not exceed $qp + q$, no great difficulty is involved in finding a limit for the degree of G within the limits of our theorem. For the present it is assumed that the degree of H_{r+1} exceeds $qp + q$. We shall return to this case later (§ 25).

13. Now H_{r+1} displaces m letters $x_1, x_2, \cdots, x_m$ which H_{rk_r} leaves fixed. These letters form one of several systems of imprimitivity of H_{r+1} which are permuted according to a primitive group. If $m = 1$, H_{r+1} is primitive. In like manner, if H_{r+i+1} displaces several letters new to H_{r+i} those letters form a system of imprimitivity, permuted with other systems according to a primitive group.

14. Let G be of the same degree, if possible, as H_{r+i}, $i = 1, 2, \cdots$. A subgroup G_1 of G displaces all but one of the letters of G and the invariant subgroup F of G_1, generated by all the substitutions of order p and degree qp in G_1, is of the same degree as G_1. Hence when H_{r+i} is contained in a primitive group of the same degree, H_{r+i} is itself primitive. F coincides with H_{r+i-1}, $i = 2, 3, \cdots$, and with H_{rk_r} when $i = 1$.

15. Consider the group in the systems of imprimitivity of m letters each ($m > 1$) of H_{r+1}. It is a primitive group and it can be shown that *it is not triply transitive*. One system, which we call s, is composed of the letters x_1, $x_2, \cdots, x_m$ and these systems of m letters can be chosen in but one way.

Now H_{rk_r} (call it F) permutes all the other systems of m letters of H_{r+1}: $t, u, \cdots$. The letters which F displaces but H_{r,k_r-1} or $H_{r-1,k_{r-1}}$ (call this group F') leaves fixed form one or several of the systems $t, u, v, \cdots$. Suppose that the group in the systems $s, t, \cdots$ of H_{r+1} is triply transitive. Then the subgroup of H_{r+1} that leaves one letter fixed is doubly transitive in the remaining systems and has an invariant subgroup generated by all its substitutions of order p and degree qp which coincides with F. Since F is an invariant subgroup of a doubly transitive group (in the systems $t, u, \cdots$), it is primitive, or an imprimitive group in which every substitution displaces all or all but one of the systems $t, u, \cdots$.* The degree of F by hypothesis exceeds qp, so that it is not regular in the systems $t, u, \cdots$. Then the subgroup of F that leaves one system fixed displaces all the systems $u, v, \cdots$, that is, all the systems of H_{r+1} except s and t; and N, the subgroup of M generated by all its substitutions of order p and degree qp, also displaces the same systems $u, v, \cdots$ as M. This is true when F is primitive in the systems in question; and when F is imprimitive, every substitution of M displaces all the systems $u, v, \cdots$, and N does not reduce to the identity because the degree of F (in the systems $t, u, \cdots$) is then the power of some prime other than p. Now F' and N coincide. For any substitution of N fixes all the letters of F that are not in F', and no other letters of F are fixed by N. It now follows since F' displaces all the letters of F except those in the systems t, that there is only the one way of dividing the letters of F into systems of m letters each.

16. Consider now the group H_{r+2} that displaces just m' letters $y_1, y_2, \cdots, y_{m'}$ new to H_{r+1}. We have seen that m' divides m and we know from the theory of primitive groups with transitive subgroups of lower degree† that m/m' is greater than 1. Call this new system of m' letters y and call the m/m' systems of H_{r+2} in the letters $x_1, x_2, \cdots, x_m$: $x', x'', \cdots$. H_{r+2} is at least doubly transitive in its systems of m' letters and contains a substitution $S = (yx') \cdots$, which certainly transforms F into itself, because $S^{-1}FS$ fixes y and x', and therefore also $x'', \cdots, x^{m/m'}$. It follows too that S permutes the systems $t, u, \cdots$ of F as units. Now consider $S^{-1}H_{r+1}S$. Its systems are $t, u, v, \cdots$, and a system s' composed of $y, x'', x''', \cdots, x^{m/m'}$. If we admit that H_{r+1} is doubly transitive in the systems, $S^{-1}H_{r+1}S$ contains a substitution $U = (s't)\cdots$. The group $H' = U^{-1}FU$ fixes all the m letters of t and the letters of x' but displaces the letters of y. Also H_{r+1} contains a substitution $V = (st)\cdots$. The group $H'' = V^{-1}H'V$ fixes all the letters of s, but displaces y. From H'' we take a substitution C of order p and degree qp that displaces y. The transitive group $S^{-1}H_{r+1}S$ contains a

*Bulletin of the American Mathematical Society, vol. 13 (1906), pp. 20–23, Theorem IV.

† These Transactions, vol. 7 (1906), pp. 499–508.

substitution W that replaces y by x'', fixes all the letters of x', and which therefore permutes the m/m' systems $y, x'', x''', \cdots, x^{m/m'}$ only among themselves. Hence $W^{-1}CW$ is a substitution similar to A_1 that fixes the letters of $y, x', x''', \cdots, x^{m/m'}$ and displaces the letters of x''. This is contrary to the hypothesis that F is the last group of the series before H_{r+1}.

17. Thus it is proved that H_{r+1} $(r > 1, m > 1)$ is not so much as triply transitive in its systems of imprimitivity of m letters each. The primitive group in the systems is simply or doubly transitive.* As we run back through the groups $F, F', \cdots$ we see that the number of new letters introduced at any step is divisible by m, and that q therefore is divisible by m; and since H_{r+1} $(r > 1)$ is not triply transitive in the systems, m is less than q.

18. Now H_{r+s} is the last imprimitive group in the series before the doubly transitive group H_{r+s+1}. Its degree is $qp + k + m + m' + \cdots + n$, where $n = m^{s-1}$, and where $qp + k$ is the degree of $F = H_{rk_r}$. We are now in position to state that *the primitive group in the systems of n letters each of H_{r+s} $(r > 1)$ is never more than doubly transitive.* It has just been proved when $m' = 1$. If m' is greater than 1, and $m'' = 1$, it is doubly and not triply transitive, because H_{r+1} is neither regular nor of class $qp + k + m - 1$.† If m'' is greater than 1, it is doubly but not triply transitive by the general theory of primitive groups with transitive subgroups of lower degree.

The J-group

19. If all the subgroups of order p and degree qp in $F = H_{rk_r}$ form a complete set of conjugates under the substitutions of F, then the largest subgroup (I_1) of H_{r+1} in which $\{A_1\}$ is invariant has a transitive constituent in the $k + m$ letters of H_{r+1} that are left fixed by A_1 (Theorem XIV). A further consequence is that all the subgroups of order p and degree qp in H_{r+1} are conjugate under the substitutions of H_{r+1}, so that in H_{r+2} the largest subgroup I_2 of H_{r+2} in which $\{A_1\}$ is invariant has a transitive constituent on the $k + m + m'$ letters of H_{r+2} that A_1 leaves fixed, and so on. Finally in H_{r+s+1}, I_{s+1} has a doubly transitive constituent of degree $k + m + m' + \cdots + n + 1$. For a transitive group is doubly transitive if it has a subgroup transitive in all but one of its letters. Let us call these transitive constituents of $I_1, I_2, \cdots, J_1, J_2, \cdots$, respectively. Then J_{s+2} is triply transitive, J_{s+3} is quadruply

*Cf. Jordan, Bulletin de la Société mathématique de France, vol. 1 (1873), pp. 185–188, where under similar conditions it is proved that the group in the systems is not quadruply transitive. It there appears to have been tacitly assumed that the group in the systems is not a simply transitive group. This is indeed the case in a corresponding passage of the discussion of primitive groups with transitive subgroups of lower degree, but in this more general problem I fail to see any valid arguments for the exclusion of those imprimitive groups whose systems of imprimitivity are permuted according to a simply transitive primitive group.

†These Transactions, loc. cit.

transitive, and so on. Similarly when $m = 1$, and H_{r+1} is a simply transitive primitive group, if all the subgroups of order p and degree qp in F are conjugate under the substitutions of F, J_1 is transitive, J_2 is at least doubly transitive, J_3 is at least triply transitive, and so on. Then if F has the required property, we can be sure that G has the same property, and that the constituent J of I, the largest subgroup of G in which $\{A_1\}$ is invariant, is transitive on all the letters of G that are left fixed by A_1.

20. If the transitive group J is of degree greater than q as at present and is alternating or symmetric, the class of G is not greater than 3, contrary to hypothesis. For if I has a substitution which displaces only letters of J, the totality of all such substitutions of I form an invariant subgroup of I, and if J is alternating or symmetric, this subgroup (assumed not to be the identity) is the alternating group. And since the largest group on the same letters in which A_1 is invariant is of order $p^q(p-1)(q!)$ (I, page 251) $\{A_1\}$ is certainly transformed into itself by substitutions leaving fixed all the letters of A_1 when J is alternating or symmetric. Hence we could apply Bochert's theorem (§ 6) to J whenever it is multiply transitive did we but know its class.

21. It is not difficult to show that the class of J is not greater than $2q-3$. *Then if J is quadruply transitive it does not displace so many as $4q-3$ letters.* While we are about it we shall prove that the class of J is not greater than $2q-4$ although this result will not be used in the proof of Theorem XIII.

Since two commutative substitutions of order p and degree qp, $q < p$, on the same letters are powers one of the other, the order of H_1 is not divisible by p^2. Then if H_2 has a transitive constituent of degree $rp+s$, $s>1$, the class of J is at most q. Let H_{i+1} be the first among the groups $H_1, H_2, \cdots$ in which any transitive constituent is of degree $rp+s$, $s>1$. Then the transitive constituents of H_i, $i>1$, are of the degrees $rp+1, r'p+1, \cdots, tp, t'p, \cdots$ ($r>1$ or $t>1$), and the order of H_i is not divisible by p^2 because its degree is less than $qp+p$. In consequence the J-group of each transitive constituent of degree $up+v$, $v>1$, in H_{i+1} is transitive of degree v, and these v letters constitute a transitive constituent of the J-group of H_{i+1}. We do not however assert that this transitive constituent of the J-group of H_{i+1} coincides with the J-group of the transitive constituent of degree $up+v$. The degree of H_{i+1} is $qp+2q-1$, $qp+2q-2$, or $qp+2q-3$, if it is argued that the class of J exceeds $2q-4$. If H_{i+1} is of degree $qp+2q-1$, H_{i+1} has a transitive constituent of degree $kp+2k-1$, k a positive integer greater than unity, and transitive constituents of degree $mp+2m$, m a positive integer or zero. The transitive constituent of degree $kp+2k-1$ is primitive and is not alternating. No such group exists if k is less than 6 (§ 2). It can be shown that the subgroup (L) of H_{i+1} that leaves one letter of the primitive constituent of degree $kp+2k-1$ fixed and which includes

H_i has just the same transitive constituents as F ($= H_{ik_i}$): $2p+2$, $p+2$, $\cdots$, $p+2$. For in the J-group of L the substitution of order 2 from F is invariant and has one invariant cycle that belongs to the constituent of F of degree $2p+2$, so that if L has a transitive constituent of degree $mp+2m$, $m>1$, the class of J is at most $2q-4$. Then the transitive constituent of degree $kp+2k-1$ in H_{i+1} is a simply transitive primitive group and in its subgroup that leaves one letter fixed the constituent of degree $2p+2$ should be simply transitive in accordance with the theorem:*

If the degree of a transitive constituent of the subgroup leaving one letter fixed in a simply transitive primitive group exceeds by two (or more) units the degree of any other transitive constituent of that subgroup, then the transitive constituent of highest degree is a simply transitive group.

But because it includes a transitive subgroup of degree $2p+1$, that constituent is doubly transitive. Let H_{i+1} be of degree $qp+2q-2$. If H_{i+1} has a primitive constituent of degree $kp+2k-2$, $k>2$, F supplies to J a substitution of degree less than $2q-2$ and of order 2. If a constituent of degree $kp+2k-2$, $k>2$, is imprimitive, systems of two letters each are permuted by that constituent according to a non-alternating primitive group, so that k is greater than 10, and F again contains substitutions which throw substitutions of order 2 and degree less than $2q-2$ into J. If H_{i+1} has one constituent of degree $2p+2$, it must have a transitive constituent of degree $mp+2m$, $m>1$, generated by similar substitutions of order p and degree mp, an imprimitive group with systems permuted according to a triply transitive group of degree $p+2$. In H_{i+1} a substitution conjugate to A_1 can be found that unites two cycles of A_1 (in letters of our imprimitive constituent) and introduces in m of its cycles exactly m new letters that form one of the $p+2$ systems of imprimitivity of the constituent in question, and that fixes the m letters of another system. Thus the class of J is something less than $2q-3$ when the degree of H_{i+1} is $qp+2q-2$. Let H_{i+1} be of degree $qp+2q-3$. Since now the J-group of H_{i+1} is of odd order, H_{i+1} has no transitive constituent of degree $mp+2m$, $m=1, 2, \cdots$, or of degree $kp+2k-2$, $k>1$. Nor has it a primitive constituent. Then H_{i+1} is an imprimitive group with systems of three letters each permuted according to a primitive group of degree $(qp+2q-3)/3$ which is not triply transitive, and which therefore does not exist if q is less than 18. The subgroup F of H_{i+1} is of degree $qp+2q-6$ and certainly has a transitive constituent of degree $rp+s$, $s>1$. Hence J is of class less than $2q-3$.

22. From this point on we confine our attention to those primitive groups G whose J-groups (whether transitive or not) contain substitutions of degree not greater than $2q-3$.

*American Journal of Mathematics, vol. 39 (1917), pp. 281–310.

Let G be of degree $qp + p$. The order of G_1, leaving one letter fixed, is not divisible by p^2 because its degree is $qp + p - 1$. It follows that the constituent J of the largest subgroup I in which $\{A_1\}$ is invariant is a transitive group of degree p. Any invariant subgroup (not identity) of J is also a transitive group of degree p. So that if I includes substitutions that leave fixed all the letters of A_1, they constitute an invariant subgroup of I and G contains a substitution of order p and degree p contrary to hypothesis. The largest group on the same letters in which $\{A_1\}$ is invariant is a subgroup of a group of order $p^q(p-1)(q!)$ in which there is just one subgroup of order p^q, generated by q cycles of order p in q $(q < p)$ different sets of letters. Now J, being of degree p, can not have an invariant subgroup of order p^2. Then J, whose order is not divisible by p^2, has an invariant subgroup of degree and order p, and is of class $p - 1$. To verify the statement that the class of J is $p - 1$ and not p, it is only necessary to notice that F of degree $qp + p - 1$ has not so many as q constituents and therefore has at least one transitive constituent of degree $rp + s$, $s > 1$. Hence $p - 1$ is less than $2q - 3$.

23. This primitive group G of degree $qp + p$ does not exist unless p is less than $2q - 2$, and then is not a subgroup of another group G' of higher degree. For if we grant that I', the largest subgroup of G' in which $\{A_1\}$ is invariant, has no substitutions on the letters of J' alone, as we must unless G' contains a transitive subgroup of degree $p + 1$, J' becomes impossible because the substitutions of order p in it should generate an invariant abelian subgroup of J' (I, Theorem VI, Corollary). But if G' contains a transitive subgroup of degree $p + 1$ the degree of G' does not exceed $2p + 2$.*

24. Let G be of degree $qp + 2p$. If the J-group of G is primitive and of degree $2p$, every invariant subgroup of J contains substitutions of order p, and I has no substitutions which fix all the letters of A_1. Therefore all the substitutions of order p in J generate an invariant abelian subgroup of J, an impossibility in a primitive group of degree $2p$.

25. Suppose now that the degree of H_{r+1} is not greater than $qp + q$. Its order therefore is not divisible by p^2 and its J_1-group is in consequence transitive. If H_{r+1} is imprimitive the degree of H_{r+s}, the last imprimitive group before the doubly transitive group H_{r+s+1}, is certainly less than $qp + q + q(1/2 + 1/4 + 1/8 + \cdots)$,† which in turn is less than $qp + 2q$. If H_{r+1} is primitive and if it permits of a quadruply transitive group H_{r+4}, the J_4 of H_{r+4} is of class not greater than q, so that by no possibility can the degree of a primitive group G which contains this primitive H_{r+1} of degree

* Marggraff, Dissertation, *Ueber primitive Gruppen mit transitiven Untergruppen geringeren Grades*, Giessen, 1889; and also, Wissenschaftliche Beilage zum Jahresberichte des Sophien-Gymnasiums zu Berlin, 1895, Programm nr. 65.

† These Transactions, vol. 7 (1906), pp. 499–508.

not greater than $qp + q$ exceed $qp + 2q + 2$. So too the class of the doubly transitive group J_1 in H_{r+s+1} (H_{r+1} imprimitive) is not greater than q, and G in that case can not be of degree greater than $qp + 2q + 2$. If p is greater than $q + 1$ none of these groups is of degree $qp + p$ or $qp + 2p$. *Then the order of none of the primitive groups G in which H_{r+1} is of degree not greater than $qp + q$ is divisible by p^2, provided q is less than $p - 1$.*

26. We have seen that if H_{r+1} is imprimitive, the last group H_{r+s}, just before H_{r+s+1}, has systems of n letters permuted according to a primitive group which is not triply transitive. Then if we assume the truth of Theorem XIII for primitive groups which contain a substitution of order p on less than q cycles, the degree of H_{r+s} is not greater than $(qp/n + 4q/n - 6)n = qp + 4q - 6n$, and the degree of the doubly transitive group H_{r+s+1} is not greater than $qp + 4q - 11$. This appears to fail when $n = q$, in which case the degree of the group in the systems is at most $p + 1$, and hence the degree of H_{r+s} is at most $qp + q$. The order of H_{r+s} is not divisible by p^2. If the degree of one of the following groups H_{r+s+t} ($t = 1, 2, \cdots$) is $qp + 2p$, the corresponding J_{s+t}-group is multiply transitive and therefore impossible. If that degree is $qp + p$, $s = t = 1$, and the group J_2 of H_{r+2} is of order $p(p - 1)$. Since J_1 is regular, the degree of H_{r+1} is at most $qp + q$ and therefore $p = q + 1$.

In the remainder of this paper H_{r+1}, wherever mentioned, is understood to be a primitive group.

27. This question must now be answered: If the order of no transitive constituent of H_{ij} is divisible by p^2, can the order of H_{ij} be divisible by p^2? Look at H_{ij} as an isomorphism between one of its transitive constituents and one other (in general intransitive) constituent. The isomorphism in question is not a direct product, for substitutions of order p must be of degree not less than qp, while there are substitutions like A_1 in H_{ij} that involve letters of both constituents and which are of degree qp. Every transitive constituent of H_{ij} is generated by its complete set of conjugate subgroups of order p. Then H_{ij} is an (m, n) isomorphism between the two constituents, and m is not divisible by p. Suppose n is divisible by p. The common quotient group is of order kp, where k is prime to p. If in the constituent of order nkp, a transitive constituent has one subgroup of order p in the subgroup of order n, every subgroup of order p of the transitive constituent, and hence every substitution of that constituent is in the subgroup of order n. But this means that all those substitutions of H_{ij} which are of order p and involve letters of the first transitive constituent of order mkp displace no letters of a certain other transitive constituent. But A_1 involves letters of all the transitive constituents. Then in the subgroup of order n of the second constituent there is no substitution of order p of any one of the transitive constituents of H_{ij}. Then n is not divisible by p, nor is the order of H_{ij} divisible by p^2.

Alternating constituents of H_{ij}

28. Let us now see if H_{ij} can have an alternating constituent.

Suppose that an alternating constituent involves letters of two or more cycles of A_1. Then we may transform A_1 by the substitution $(a_1\, a_2\, a_3) \cdots$ of H_{ij} into a substitution C which has $(a_2\, a_3\, a_1\, a_4\, a_5 \cdots a_p)$ for its first cycle, $(b_1\, b_2\, b_3 \cdots b_p)$ for its second cycle. If C has two letters new to

$$A_1 = (a_1\, a_2\, a_3 \cdots a_p)(b_1\, b_2\, b_3 \cdots b_p) \cdots$$

in any cycle, in some power C^n, $n = 1, 2, \cdots$, or $(p-1)/2$, these two new letters are adjacent and in $C^{-n} A_1 C^n$ the second of the two new letters is omitted. Now if $\{A_1, C^{-n} A_1 C^n\}$ is again an alternating group in the constituent on the letters $a_1, a_2, \cdots, a_p$ the process may be repeated until we have C_1, similar to A_1, with $(a_2\, a_3\, a_1\, a_4 \cdots a_p)$ and $(b_1\, b_2\, b_3 \cdots b_p)$ for its first two cycles. Write $A = (123 \cdots p)$, and $B = (132)A(123)$. Then

$$\begin{aligned}
B^{-n} AB^n &= (132)A^{-n}(123)A(132)A^n(123) \\
&= (132)A^{-n}A(234)(132)A^n(123) \\
&= A(243)A^{-n}(134)A^n(123) \\
&= A(243)(1+n,\ 3+n,\ 4+n)(123) \\
&= A(12534) && (\text{for } n = 1), \\
&= A(12456) && (\text{for } n = 2), \\
&= A(12674) && (\text{for } n = 3),
\end{aligned}$$

and

$$= A(124)(1+n,\ 3+n,\ 4+n) \quad (3 < n \leqq (p-1)/2.)$$

Now the group $\{A, B^{-n} AB^n\}$ is a primitive group of degree p and of class at most 6, and hence is alternating. It is now evident that $\{A_1, C_1\}$ of degree less than $qp + q$ has two simple constituents of degree p each, one alternating and one cyclic, so that G certainly contains a substitution of order p and of degree less than qp. Then no alternating constituent of H_{ij} involves letters of two or more cycles of A_1.

29. Suppose an imprimitive constituent of H_{ij} has systems of imprimitivity permuted according to an alternating group such that A_1 permutes more than p of these systems. Then we may read the preceding paragraph with the understanding that $a_1, a_2, \cdots, b_1, b_2, \cdots$, are not single letters but are systems of imprimitivity, and draw the conclusion that A_1 does not permute more than p of these systems.

30. Any alternating constituent of H_{ij} implies the existence of (1) a subgroup $\{A_1, C\}$, where C is the transform of A_1 by the substitution $(a_1\, a_2\, a_3)(a_4) \cdots$ of H_{ij}, and (2) a subgroup $E_1 = \{A_1, C_1\}$ of the latter in which C_1,

a transform of A_1, has at most one letter new to A_1 to a cycle, and generates with A_1 an alternating group in the p letters $a_1, a_2, \cdots, a_p$.

When the prime p is greater than 7 a transitive group simply isomorphic to an alternating group of degree p is, if of degree greater than p, of degree $(p-1)p/2$. Corresponding to values of p greater than 7 there is this one transitive representation of the alternating group and no other of degree less than $(q-1)p+q$. For the two largest intransitive subgroups of the alternating group of degree p are of the orders $(p-2)!$ and $3(p-3)!$; the largest imprimitive subgroup is of order $[(p-1)/2]!\,2!$; the two largest alternating subgroups of the alternating group of degree p are of the orders $(p-1)!/2$ and $(p-2)!/2$; the largest non-alternating primitive subgroup of the alternating p-group is of index greater than $p(p-1)$.

31. Now for the first time we shall use the condition that p is greater than $2q-5$.

Since p is greater now than $2q-5$ this transitive constituent of degree $(p-1)p/2$ simply isomorphic to the alternating group of degree p can occur only if E_1 has at most three transitive constituents and is of degree at most $qp+2$. Whence it follows that H_{r+1} is of degree not greater than $qp+2q+2$.

32. If all the transitive constituents of E_1 are alternating groups, E_1 contains a substitution of any odd prime order P less than p and of degree qP. In particular there is a prime P such that $(q+1)/2 < P \leqq q-1$;* hence one of the two numbers, $qP+q^2-q$ or $2q^2-P^2$, either of which is less than $qp+q$, can be used for the limit of the degree of H_{r+1}, at least if H_{r+1} is simply transitive.

If H_{r+1} is doubly transitive it can contain no transitive subgroup of lower degree generated by substitutions that have fewer than p cycles each. For if H_{r+1} had such a transitive subgroup, its subgroup that leaves one letter fixed, and in which F is invariant, would also be an imprimitive group with no systems of so many as p letters. But F is intransitive.

33. Let there be at least one transitive constituent of order greater than $p!/2$ in multiple isomorphism to the alternating constituent of degree p. From the manner of the derivation of C_1, we know that C_1 can be taken similar to A_1 in the two generators of any transitive constituent of E_1. In a primitive constituent of order greater than $p!/2$ the head that corresponds to the identity of the alternating constituent is of order $(mp+n)k$ $(n \leqq m < q)$. If $n=0$, E_1 contains a substitution of order p and of degree less than qp. Then let n be greater than zero. The subgroup of this constituent that leaves one letter fixed is of order $k(p!/2)$ and has an invariant subgroup of

* Tchebychef, Journal de mathématiques, vol. 17 (1852), pp. 366–390; *Oeuvres*, 1899, vol. 1, pp. 51–70.

order k (which may be unity), with respect to which the quotient group is an alternating group of order $p!/2$. The subgroup of E_1 that leaves this letter fixed has the same alternating constituent as before. Call it E'; if E' has a primitive constituent of order greater than $p!/2$, pass on to E''; we must at last reach a group without a primitive constituent of order greater than $p!/2$ or with such a constituent of degree mp. Hence we may as well assume in the first place that E_1 has no primitive constituent of order greater than $p!/2$.

34. Let there be an imprimitive constituent of order greater than $p!/2$. If the invariant subgroup that corresponds to the identity of the alternating constituent is transitive, we have the same condition as when it was assumed primitive. Then the head may be taken to be intransitive. The two similar generators of order p permute systems of less than q letters. The largest possible systems are permuted according to a primitive group. Now the transitive sets of the head are systems of imprimitivity. There is no larger head because the quotient group with respect to the first is alternating. The group according to which these systems are permuted is simply isomorphic to the alternating group, that is, is the alternating group of degree p or its transitive representation on $p(p-1)/2$ letters, which last is impossible when q is less than p and p is greater than 7. Then if there are m letters in each system of imprimitivity of a certain transitive constituent, the degree of this imprimitive constituent is mp. Let P be the largest prime less than $2q-6$ so that an imprimitive constituent of E_1 has a substitution of order P which permutes systems, and since m, less than $q-2$, is less than P, it displaces no other letter of that constituent. Hence E_1 has a substitution of order P and degree qP. The substitutions of order P and degree qP in E_1 generate an invariant subgroup of E_1 which is an alternating-p group or has an alternating group of degree p as a quotient group. But this subgroup coincides with E_1 because it contains all E_1's substitutions of order p, two of which generate E_1. Take, in one of the imprimitive constituents of E_1, a subgroup that leaves $p-P$ systems fixed. This subgroup (E) is an isomorphism between alternating groups of degree P and imprimitive groups (for it respects the systems of E_1) having alternating quotient groups. Its degree is qP. Not all the constituents are alternating groups (we have discussed such a possibility before). Since H_{r+1} is primitive the substitutions of order P and degree qP in H_{r+1} generate a transitive group. Imprimitive constituents of E are of the degrees M_1P, M_2P, $\cdots$, $(M_1 \geqq M_2 \geqq \cdots)$. If M_1 is greater than $P/2$, E has at most $q-M_1+1$ transitive constituents. Therefore H_{r+1} contains a transitive subgroup of degree not greater than

$$qP + (q-M_1)q \leqq q^2 + qP/2 - q/2 < 2q^2 - 3q < qp.$$

If M_1 is less than $P/2$, there are in E substitutions of order P' and degree qP',

where P' is the largest prime less than P, and certain ones of which, just as before, generate a group E' with imprimitive constituents of degree $M_1' P'$, $M_2' P'$, $\cdots$ $(M_1' \geqq M_2' \geqq \cdots)$, each of which are alternating-P' groups in their systems of imprimitivity. Not all the transitive constituents are alternating groups of degree P', nor is M_1' greater than $P'/2$. Another subgroup E'' generated by substitutions of prime order P'' and of degree qP'', where P'' is the largest prime less than P', and with imprimitive constituents of degree $M_1'' P''$, $M_2'' P''$, $\cdots$ $(M_1'' \geqq M_2'' \geqq \cdots)$ can next be set up, and so on. Thus we can find a substitution of order p' and degree qp', $q - 1 \geqq p' > (q+1)/2$, which insures that the degree of H_{r+1} is at most

$$qp' + q^2 - q < 2q^2 - 2q \leqq qp + q,$$

or else is at most

$$2q^2 - p'^2 < 2q^2 - (q+1)^2/4 < qp + q.$$

With the next paragraph in view, it should be noticed that E_1 may be free from alternating constituents of degree p without affecting the above conclusion.

35. Any imprimitive constituent of H_{ij} is generated by substitutions of order p, one of which must permute systems, so that the number of letters in a system of imprimitivity is at most q and all these generators permute systems. There are at least p systems in such an imprimitive constituent. Suppose that the group in the systems is alternating. Then H_{ij} contains a substitution which in that constituent is a circular permutation of just three systems. Transform A_1 by it, and the transform C generates with A_1 a group with an imprimitive constituent of degree $m'p$, which is alternating in its systems. Continuing step by step we can find a group $E_1 = \{A_1, C_1\}$, in which C_1, similar to A_1, generates with A_1 an imprimitive constituent, alternating in the systems, and has not more than one new letter in a cycle. All the transitive constituents of E_1 are imprimitive groups with systems permuted according to alternating groups of degree p.

Hence H_{ij}, when p is greater than $2q - 5$, can have no transitive constituent which is alternating or permutes systems of imprimitivity according to an alternating group.

The degree of F

36. If H_{r+1} is a primitive group, its subgroup F $(= H_{rk_r})$, of degree lower by unity, is an isomorphism between groups generated by substitutions of order p, which are not alternating, nor, if imprimitive, are the groups in the systems alternating. Then if we assume the truth of our theorem for smaller values of q than the one actually under consideration, we may say that F is an isomorphism between x_i primitive constituents of degree at most $q_i p/x_i + 4q/x_i - 4$ (or by the same symbolism an imprimitive constituent of degree

$q_i(p+4)-4x_i$ with systems of x_i letters each permuted according to a primitive group of degree $q_i\,p/x_i+4q_i/x_i-4)$, $i=1,2,\cdots$, and y constituents of degree p, y_1 of degree $p+1$, y_2 of degree $p+2$, including imprimitive groups with systems permuted according to groups of these degrees, making the degree of F at most

$$yp+y_1(p+1)+y_2(p+2)+\sum x_i(q_i\,p/x_i+4q_i/x_i-4)$$
$$=qp+4q-4y-3y_1-2y_2-4\sum x_i.$$

Here $\sum x_i$ is not zero, so that the maximum degree of F is got by putting $\sum x_i=1$, $y_2=1$, and $y=y_1=0$. This maximum degree is then $qp+4q-6$, so that the maximum degree of H_{r+1} appears to be $qp+4q-5$. The next highest degree of F is got by putting $\sum x_i=1$, $y=0$, $y_1=1$, and $y_2=0$, whence F is of degree $qp+4q-7$; and after this we have the degree $qp+4q-8$, by putting $\sum x_i=2$, $y=y_1=y_2=0$, or by putting $\sum x_i=1$, $y=y_1=0$, $y_2=2$, etc. This last limit, $qp+4q-8$ we may accept, for when H_{r+1} is a simply transitive primitive group a constituent of F on more than half the letters of F is not more than simply transitive (by the theorem quoted in § 21). Then when there are just two transitive constituents in F, and one is of degree $p+1$ or $p+2$, the other constituent is of degree not greater than $(q-1)p+4(q-1)-7$, making the degree of F at most $qp+4q-8$. If the large constituent is imprimitive, we have $\sum x_i$ greater than unity. If H_{r+2}, H_{r+3}, H_{r+4} exist, their degrees are not greater than $qp+4q-6$, $qp+4q-5$, $qp+4q-4$, respectively. A group H_{r+5} of degree $qp+4q-3$ that contains the primitive group H_{r+1} we have seen to be impossible because its group J_5 is a quintuply transitive group of class less than $2q-2$, the degree of which therefore can not be so great as $4q-3$ (§ 21).

37. If p is greater than $2q-5$, $2p$ is greater than $4q-7$, so that the degree of H_{r+1} can not be $qp+2p$. It was noted (§ 24) that no one of the groups H_{r+2}, H_{r+3}, $\cdots$ is of degree $qp+2p$, and if at last we impose the strong condition, that p be greater than $2q-3$, the degree of no one of our primitive groups is $qp+p$ (§ 23). Then the order of G is in no case divisible by p^2, and the proof of Theorem XIII by induction is complete.

38. The degrees of the primitive groups of class 4 and of class 6 do not exceed 8 and 10 respectively so that it is possible to formulate the clean-cut

THEOREM XV. *The degree of a primitive group of class greater than* 3 *which contains a substitution of prime order* p *on* q *cycles* ($p>2q-3$, $q>1$) *does not exceed* $qp+4q-4$.

STANFORD UNIVERSITY,
December, 1916

ON THE DEGREE OF CONVERGENCE OF BIRKHOFF'S SERIES*

BY

W. E. MILNE

INTRODUCTION

In the linear differential equation

$$\frac{d^n u}{dx^n} + P_2(x)\frac{d^{n-2} u}{dx^{n-2}} + \cdots + P_n(x)u + \lambda u = 0 \tag{1}$$

let the coefficients $P_2(x), \cdots, P_n(x)$ be continuous with their derivatives of all orders in the closed interval (a, b), and let

$$W_1(u) = 0, \quad W_2(u) = 0, \quad \cdots, \quad W_n(u) = 0 \tag{2}$$

be n linearly independent linear homogeneous conditions in $u(a), u'(a), \cdots, u^{(n-1)}(a); u(b), u'(b), \cdots, u^{(n-1)}(b)$, which have been *normalized* and are *regular*.† Then it is known that the characteristic numbers of this system are in general simple, are infinite in number, and have no cluster point in the finite plane. If $U_1(x), U_2(x), U_3(x), \cdots$ denote solutions of the system corresponding to simple characteristic numbers $\lambda_1, \lambda_2, \lambda_3, \cdots$, then an arbitrary function $f(x)$ may be represented, formally at least, by a series

$$A_1 U_1(x) + A_2 U_2(x) + A_3 U_3(x) + \cdots. \tag{3}$$

The term $A_i U_i(x)$ is given by the equation

$$A_i U_i(x) = \int_a^b f(t) R_i(x, t, \lambda)\, dt, \tag{4}$$

where $R_i(x, t, \lambda)$ is the residue at $\lambda = \lambda_i$ of the Green's function $G(x, t, \lambda)$ belonging to the system (1) and (2). If the characteristic numbers are not all simple we shall agree to define the terms of the series (3) by means of equation (4). Therefore in any case the sum of the first ν terms of (3) is

$$S_\nu(x) = \frac{1}{2\pi\sqrt{-1}} \int_{\Gamma_\nu} \int_a^b f(t) G(x, t, \lambda)\, dt d\lambda, \tag{5}$$

* Presented to the Society, December 28, 1915. The general nature of the series treated here was first discussed by Birkhoff, these Transactions, vol. 9 (1908), pp. 373–395. This paper is based on his results.

† For definitions of these terms see Birkhoff, loc. cit., p. 382.

where Γ_ν is a contour in the λ-plane inclosing the first ν characteristic numbers, and excluding all others.

It is the purpose of this paper to discuss the degree of convergence of the series (3) with suitable restrictions on the function $f(x)$. This is done by considering the order of magnitude of the remainder $f(x) - S_\nu(x)$. The methods are similar on the one hand to those used by Birkhoff in proving the convergence of (3), and on the other hand to those used by Jackson in finding the degree of convergence of Sturm-Liouville series.* It is found that if $f(x)$ with its first $m-1$ derivatives vanish at a and at b, the remainder after ν terms of the series is less than a constant multiple of $1/\nu^m$ when the mth derivative is of limited variation, and less than a constant multiple of $(\log \nu)/\nu^{m+1}$ when the mth derivative satisfies a Lipschitz condition.

1. Preliminary formulas

In this section the needed facts concerning the system (1) and (2) are briefly presented.

1. **The characteristic numbers.** The characteristic numbers of the system (1) and (2) form a pair of infinite sequences

$$(6)\qquad \begin{aligned} \lambda_\nu' &= -\left(\frac{2\nu\pi\sqrt{-1}}{b-a}\right)^n (1 + E'/\nu), \\ \lambda_\nu'' &= -\left(\frac{-2\nu\pi\sqrt{-1}}{b-a}\right)^n (1 + E''/\nu) \quad (\nu = 1, 2, 3, \cdots), \end{aligned}$$

where E' and E'' are functions of ν that are bounded as ν becomes infinite.

2. **The Green's function.** The explicit form of the Green's function is

$$(7)\quad G(x,t,\lambda) = \frac{(-1)^n}{\Delta}\begin{vmatrix} y_1(x) & y_2(x) & \cdots & y_n(x) & \overline{G}(x,t,\lambda) \\ W_1(y_1) & W_1(y_2) & \cdots & W_1(y_n) & W_1(\overline{G}) \\ W_2(y_1) & W_2(y_2) & \cdots & W_2(y_n) & W_2(\overline{G}) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ W_n(y_1) & W_n(y_2) & \cdots & W_n(y_n) & W_n(\overline{G}) \end{vmatrix},$$

where

$$(8)\qquad \Delta = \begin{vmatrix} W_1(y_1) & W_1(y_2) & \cdots & W_1(y_n) \\ W_2(y_1) & W_2(y_2) & \cdots & W_2(y_n) \\ \cdots & \cdots & \cdots & \cdots \\ W_n(y_1) & W_n(y_2) & \cdots & W_n(y_n) \end{vmatrix},$$

and

* These Transactions, vol. 15 (1914), pp. 439–466.

$$(9)\quad \bar{G}(x,t,\lambda) = \pm\frac{1}{2}\sum_{i=1}^{n} y_i(x)\,\bar{y}_i(t)$$

$$= \pm\frac{1}{2}\begin{vmatrix} y_1(x) & y_2(x) & \cdots & y_n(x) \\ y_1^{(n-2)}(t) & y_2^{(n-2)}(t) & \cdots & y_n^{(n-2)}(t) \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ y_1(t) & y_2(t) & \cdots & y_n(t) \end{vmatrix} \Big/ \begin{vmatrix} y_1^{(n-1)}(t) & y_2^{(n-1)}(t) & \cdots & y_n^{(n-1)}(t) \\ y_1^{(n-2)}(t) & y_2^{(n-2)}(t) & \cdots & y_n^{(n-2)}(t) \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ y_1(t) & y_2(t) & \cdots & y_n(t) \end{vmatrix}$$

$$(+ \text{ if } x > t,\ - \text{ if } x < t).$$

Here $y_1(x), y_2(x), \cdots, y_n(x)$ are any n linearly independent solutions of equation (1).

3. **Asymptotic form of solutions.** Let $\lambda = \rho^n$, let the ρ-plane be divided into $2n$ equal sectors

$$J_k: \qquad k\pi/n \leqq \arg\rho \leqq (k+1)\pi/n \quad (k = 0, 1, \cdots, n-1),$$

and let the n nth roots of $-1, w_1, w_2, \cdots, w_n$, be numbered in such a manner that when ρ is on any given sector J_k the inequalities

$$R(\rho w_1) \leqq R(\rho w_2) \leqq \cdots \leqq R(\rho w_n)$$

are satisfied, where $R(\rho w_i)$ denotes the real part of ρw_i. Then if the coefficients $P_s(x)$ in (1) have continuous derivatives of order $(m+n-s)$ in (a,b), m being any positive integer or zero, there exist, for every sector J_k, n independent solutions of (1) of the form

$$(10)\qquad y_i(x) = \bar{u}_i(x,\rho) + e^{\rho w_i(x-c_i)}E_{i0}/\rho^{m+1},$$

$$y_i^{(j)}(x) = \bar{u}_i^{(j)}(x,\rho) + e^{\rho w_i(x-c_i)}E_{ij}/\rho^{m-j+1} \quad (j = 1, 2, \cdots, n-1),$$

in which

$$(11)\quad \bar{u}_i(x,\rho) = e^{\rho w_i(x-c_i)}[1 + \phi_1(x)/\rho w_i + \cdots + \phi_m(x)/(\rho w_i)^m].$$

The functions $\phi_j(x)$ are independent of i and ρ, and have continuous derivatives of order $(m+n-j)$ in (a,b), and the E_{ij} are analytic in ρ and bounded as ρ becomes infinite in J_k for all values of x in (a,b). Likewise the n functions $\bar{y}_i(t)$ of formula (9) have the form

$$(12)\quad \bar{y}_i(t) = \frac{e^{-\rho w_i(t-c_i)}}{n(\rho w_i)^{n-1}}[1 + \psi_1(t)/\rho w_i + \cdots + \psi_m(t)/(\rho w_i)^m + \bar{E}_i/\rho^{m+1}],$$

in which the $\psi_j(t)$ and the $\bar{E}_i$ have the same properties as the $\phi_j(x)$ and the E_{ij} respectively.*

4. **The constants c_i, and the functions v_i and u_i.** The constants c_i in each of the functions y_i and $\bar{y}_i$ may be chosen as follows:

$$c_i = a \text{ if } R(\rho w_i) < 0, \qquad c_i = b \text{ if } R(\rho w_i) > 0.$$

When n is odd and $i = (n+1)/2$, c_i has different values in the two halves of the sector J_k, but in all other cases c_i remains constant throughout the sector. By the given choice of the c_i we see that when x is in (a, b),

$$|e^{\rho w_i(x-c_i)}| \leqq 1 \qquad (i = 1, 2, \cdots, n).$$

Also if we let

$$c_i' = b \text{ if } R(\rho w_i) < 0, \qquad \text{and} \qquad c_i' = a \text{ if } R(\rho w_i) > 0,$$

then for t in (a, b),

$$|e^{-\rho w_i(t-c_i')}| \leqq 1 \qquad (i = 1, 2, \cdots, n).$$

It will be convenient to introduce the notation

$$v_i(t) = n(\rho w_i)^{n-1}\bar{y}_i(t) \qquad (i = 1, 2, \cdots, n), \tag{13}$$

and

$$u_i(t) = n(\rho w_i)^{n-1} e^{\pm\rho w_i(b-a)}\bar{y}_i(t) \tag{14}$$

$$(+ \text{ if } R(\rho w_i) < 0, \ - \text{ if } R(\rho w_i) > 0).$$

The explicit form of $u_i(t)$ is

$$u_i(t) = e^{-\rho w_i(t-c_i')}\left[1 + \frac{\psi_1(t)}{\rho w_i} + \cdots + \frac{\psi_m(t)}{(\rho w_i)^m} + \frac{\bar{E}_i}{\rho^{m+1}}\right]. \tag{15}$$

5. **The boundary conditions.** The normalized conditions (2) have the form†

$$W_i(u) = A_i(u) + B_i(u) = 0,$$

where

$$A_i(u) = \alpha_i u^{(k_i)}(a) + \sum_{j=0}^{k_i-1}\alpha_{ij}u^{(j)}(a),$$

$$B_i(u) = \beta_i u^{(\lambda_i)}(b) + \sum_{j=0}^{k_i-1}\beta_{ij}u^{(j)}(b) \quad (k_1 \geqq k_2 \geqq \cdots \geqq k_n),$$

and no three successive k's are equal.

When the values of y_i as given in (10) are substituted for u in the above expressions it is found that

$$W_j(y_i) = \rho^{k_j}U_{ij} \qquad (i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, n), \tag{16}$$

* The formulas (10), (11), and (12) were first obtained by Birkhoff in a less explicit form, these Transactions, vol. 9 (1908), pp. 219–231. The forms used here are given in a note by the author in the Proceedings of the National Academy of Sciences, vol. 2 (1916), p. 543, and in the Bulletin of the American Mathematical Society, vol. 23 (1917), pp. 166–169.

† See second footnote on the first page of the article.

where U_{ij} is bounded as ρ becomes infinite on the given sector. Also when $R(\rho w_i) < 0$,

$$B_j(y_i) = \rho^{k_j} e^{\rho w_i (b-a)} V_{ij}, \tag{17}$$

and when $R(\rho w_i) > 0$,

$$A_j(y_i) = \rho^{k_j} e^{-\rho w_i (b-a)} V_{ij} \qquad (j = 1, 2, \cdots, n), \tag{18}$$

where V_{ij} is bounded on J_k as ρ becomes infinite.

6. **Continuity of derivatives of the coefficients.** Birkhoff assumed throughout his proof of convergence that the functions $P_s(x)$ had derivatives of all orders, but of this property he made no other use than to establish the existence of solutions like those in (10). Since such solutions exist if the functions $P_s(x)$ have continuous derivatives of order $m + n - s$, where m is any positive integer or zero, it is evident that all the properties of the system (1) and (2) that we have mentioned will hold with this modification of the hypotheses. Therefore it will no longer be assumed that $P_s(x)$ has derivatives of all orders.

2. Transformation of the contour integral

In this section the contour integral in (5) is put in a more serviceable form.

From (7), (9), and (13) it follows that after the change of variable $\lambda = \rho^n$, the contour integral can be put in the form

$$\sum_{i=1}^{n} \frac{(-1)^{n-1} w_i}{4\pi\sqrt{-1}} \int_{\gamma_\nu} \int_a^b \frac{f(t)}{\Delta} \cdot \begin{vmatrix} y_1(x) & y_2(x) & \cdots & y_n(x) & \pm y_i(x) \\ W_1(y_1) & W_1(y_2) & \cdots & W_1(y_n) & \overline{W}_1(y_i) \\ W_2(y_1) & W_2(y_2) & \cdots & W_2(y_n) & \overline{W}_2(y_i) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ W_n(y_1) & W_n(y_2) & \cdots & W_n(y_n) & \overline{W}_n(y_i) \end{vmatrix} v_i(t)\, dt\, d\rho \tag{19}$$

$$(+ \text{ if } x > t, \; - \text{ if } x < t),$$

where $\overline{W}_j(u) = -A_j(u) + B_j(u)$. The path of integration γ_ν is the segment $\gamma_{\nu,k}$ of a circle $|\rho| = r_\nu$ on the sector J_k, joined to the segment $\gamma_{\nu,k+1}$ of the same circle on the adjacent sector J_{k+1}.*

In (19) when $R(\rho w_i) > 0$ subtract the ith column from the last column, but when $R(\rho w_i) < 0$ add the ith column to the last column. Take the term in the upper right-hand corner out of the determinant and write it separately, putting zero in its place. Replace $W_j(y_i)$ and $A_j(y_i)$ and $B_j(y_i)$ (the last two of which will appear in the last column after the addition or subtraction) by their values from (16), (17), and (18). Cancel the factor

* When ρ is on J_k the functions y_i and $\overline{y}_i$ must be those belonging to J_k, and when ρ is on J_{k+1} they must be those belonging to J_{k+1}.

ρ^{k_j} from the $(j+1)$th row of the determinant in (19) with the same factor in the jth row of Δ. And finally remove the factor $e^{\mp\rho w_i(b-a)}$ ($+$ if $R(\rho w_i) < 0$, $-$ if $R(\rho w_i) > 0$) from the last column and combine it with $v_i(t)$, giving $u_i(t)$. When these steps have been taken it will be found that (19) takes the form

$$(20)\quad -\sum_{i=1}^{n}\frac{w_i}{2\pi\sqrt{-1}}\int_{\gamma_\nu}\int_{c_i}^{x} f(t)\,y_i(x)\,v_i(t)\,dt\,d\rho$$

$$-(-1)^n\sum_{i=1}^{n}\frac{w_i}{2\pi\sqrt{-1}}\int_{\gamma_\nu}\frac{\begin{vmatrix} y_1(x) & y_2(x) & \cdots & y_n(x) & 0 \\ U_{11} & U_{12} & \cdots & U_{1n} & V_{1i} \\ U_{21} & U_{22} & \cdots & U_{2n} & V_{2i} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ U_{n1} & U_{n2} & \cdots & U_{nn} & V_{ni} \end{vmatrix}}{\begin{vmatrix} U_{11} & U_{12} & \cdots & U_{1n} \\ U_{21} & U_{22} & \cdots & U_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ U_{n1} & U_{n2} & \cdots & U_{nn} \end{vmatrix}}\cdot\left[\int_a^b f(t)\,u_i(t)\,dt\right]d\rho.$$

The sequence of paths of integration γ_ν may be so chosen that the distance from the nearest ρ_i corresponding to a characteristic number λ_i is uniformly greater than some constant δ. Then when ρ is on any path γ_ν the determinant in the denominator of (20) remains uniformly greater in absolute value than some positive constant d. It is therefore evident that after the determinant in the numerator is expanded by minors of the top row, the second sum in (20) may be written

$$(21)\qquad R_\nu(x) = \sum_{i=1}^{n}\sum_{j=1}^{n}\int_{\gamma_\nu} E_{ij}\,y_j(x)\left[\int_a^b f(t)\,u_i(t)\,dt\right]d\rho,$$

where the E_{ij} are bounded as ρ becomes infinite along the sequence of paths γ_ν.

From (10), (12), and (13) it follows that*

$$(22)\qquad y_i(x)\,v_i(t) = e^{\rho w_i(x-t)}\left[1+\sum_{j=1}^{m}\omega_j(x,t)/(\rho w_i)^j + E_i/\rho^{m+1}\right],$$

where $\omega_j(x,t)$ is equal to

$$\sum_{k=1}^{j-1}\phi_k(x)\,\psi_{j-k}(t),$$

and therefore has a continuous $(m+n-j)$th derivative with respect to t in (a,b).

* The letter E will regularly be used to denote functions of ρ and x that are bounded on the sector to which they belong as ρ becomes infinite.

Let

$$\xi(x, t, \rho w_i) = e^{\rho w_i(x-t)} \sum_{j=1}^{m} \omega_j(x, t)/(\rho w_i)^j.$$

Then if we make the transformation $\rho w_i = z$ in the ith term of the sum

$$\frac{-1}{2\pi\sqrt{-1}} \sum_{i=1}^{n} \int_{\gamma_\nu} \int_{c_i}^{x} f(t)\xi(x, t, \rho w_i)\, dt\, w_i\, d\rho \quad (i = 1, 2, \cdots, n),$$

and combine the new paths of integration, the result may be written

$$L_\nu(x) = \frac{-1}{2\pi\sqrt{-1}} \left[\int_{l_1} \int_{a}^{x} f(t)\xi(x, t, z)\, dt\, dz - \int_{l_2} \int_{x}^{b} f(t)\xi(x, t, z)\, dt\, dz \right], \tag{23}$$

where l_1 is the semicircle of radius $|z| = r_\nu$ in the left half of the z-plane, and l_2 is the semicircle in the right-half plane.*

Similarly

$$\frac{-1}{2\pi\sqrt{-1}} \sum_{i=1}^{n} \int_{\gamma_\nu} \int_{c_i}^{x} f(t)\, e^{\rho w_i(x-t)}\, dt\, w_i\, d\rho = \frac{-1}{2\pi\sqrt{-1}} \left[\int_{l_1} \int_{a}^{x} f(t)\, e^{z(x-t)}\, dt\, dz - \int_{l_2} \int_{x}^{b} f(t)\, e^{z(x-t)}\, dt\, dz \right]. \tag{24}$$

The right-hand side of this last equation we can simplify further by changing the order of integration and noting that

$$-\int_{l_1} e^{z(x-t)}\, dz = \int_{l_2} e^{z(x-t)}\, dz.$$

For then the right-hand side of (24) becomes

$$\frac{1}{2\pi\sqrt{-1}} \int_{a}^{b} f(t) \int_{-\sqrt{-1}r_\nu}^{\sqrt{-1}r_\nu} e^{z(x-t)}\, dz\, dt,$$

which upon integration with respect to z becomes

$$\frac{1}{\pi} \int_{a}^{b} f(t) \frac{\sin r_\nu(x - t)}{x - t}\, dt.$$

It remains to treat the terms arising from the integration of the E terms in (22). We restrict $f(x)$ to be bounded in (a, b), and let $M = \max|f(x)|$ in (a, b). Since E_i is bounded as ρ becomes infinite along the sequence of paths γ_ν for all values of x and t in (a, b), we may let $K = \max|E_i|$ $(i = 1, 2, \cdots, n)$. Now set $\rho w_i = r_\nu(\sin\phi + \sqrt{-1}\cos\phi)$, let g_i be the arc on

*A similar transformation is used by Tamarkine, Rendiconti del Circolo Matematico di Palermo, vol. 34 (1912), pp. 345–382.

which ρw_i lies when ρ is on γ_ν, and let ϕ_1 and ϕ_2 be the angles made with the axis of imaginaries by radii drawn to the extremities of g_i. Let $|x - c_i| = h$. Then

$$\left| \int_{\gamma_\nu} \int_{c_i}^{x} f(t) \frac{e^{\rho w_i (x-t)} E_i}{\rho^{m+1}} dt\, w_i\, d\rho \right|$$

$$\leqq \frac{MK}{r_\nu^m} \left| \int_{\phi_1}^{\phi_2} \int_{c_i}^{x} e^{r_\nu \sin\phi (x-t)} dt\, d\phi \right| = \frac{MK}{r_\nu^{m+1}} \left| \int_{\phi_1}^{\phi_2} \frac{1 - e^{-r_\nu h \sin\phi}}{\sin\phi} d\phi \right|$$

$$\leqq \frac{2MK}{r_\nu^{m+1}} \int_0^{\pi/2} \frac{1 - e^{-r_\nu h \sin\phi}}{\sin\phi} d\phi < \frac{2MK}{r_\nu^{m+1}} \left[h r_\nu \int_0^{1/r_\nu} d\phi + \int_{1/r_\nu}^{\pi/2} \frac{\pi d\phi}{2\phi} \right]$$

$$= \frac{2MK}{r_\nu^{m+1}} \left[h + \frac{\pi}{2} \log (\pi r_\nu / 2) \right] = M \frac{E_\nu \log r_\nu}{r_\nu^{m+1}},$$

where E_ν is independent of $f(x)$ and is bounded as r_ν becomes infinite. To verify the last inequality we have only to note that

$$\frac{1 - e^{-r_\nu h \sin\phi}}{\sin\phi}$$

is at most equal to $r_\nu h$ in the interval $0 \leqq \phi \leqq 1/r_\nu$, and is less than $\pi/(2\phi)$ in the interval $1/r_\nu \leqq \phi \leqq \pi/2$.

Let $\phi(\nu)$ and $\psi(\nu)$ be any two real functions of ν. Then if $\phi(\nu)/\psi(\nu)$ is bounded as ν becomes infinite, that fact will be indicated by the notation

$$\phi(\nu) = O(\psi(\nu)).$$

In this notation

$$\frac{E_\nu \log r_\nu}{r_\nu^{m+1}} = O\left(\frac{\log \nu}{\nu^{m+1}} \right),$$

as is readily seen from (6), since $2r_\nu = \sqrt[n]{|\lambda_\nu|} + \eta_\nu$, where η_ν may be taken less than some constant independent of ν.

Collecting the results obtained so far, we write the sum of ν terms of the series as follows:

$$(25) \quad S_\nu(x) = \frac{1}{\pi} \int_a^b f(t) \frac{\sin r_\nu (x-t)}{x-t} dt + L_\nu(x) + R_\nu(x) + O\left(\frac{\log \nu}{\nu^{m+1}} \right).$$

The constant implied in the O-notation is independent of x, so the equation (25) holds *uniformly* with respect to x.

3. Order of magnitude of the terms $L_\nu(x)$ and $R_\nu(x)$

In this section the terms $L_\nu(x)$ and $R_\nu(x)$ are evaluated for a restricted class of functions $f(x)$. The conclusions are given in Lemmas 2 and 3, for the proof of which the following lemma is needed.

Lemma 1. *If $F(x)$ is any continuous function of limited variation in (a, b),*

then

$$\int_a^b F(t)\, e^{-\rho w_i(t-c_i')}\, dt = E/r_\nu \qquad (i = 1, 2, \cdots, n),$$

where as usual $r_\nu = |\rho|$.

Since $F(t)$ is continuous and of limited variation in (a, b), there exist two continuous functions F_1 and F_2, each positive and monotone increasing in (a, b), and such that $F = F_1 - F_2$. Write $\rho w_i = r_\nu(p + \sqrt{-1}\, q)$, where p and q are real. Then in the case where $R(\rho w_i) < 0$ and where consequently $c_i' = b$, we have

$$\int_a^b F(t)\, e^{-\rho w_i(t-b)}\, dt$$
$$= \sum_{j=1}^{2} (-1)^{j-1} \int_a^b F_j(t)\, e^{-r_\nu p(t-b)} [\cos r_\nu q(t-b) - \sqrt{-1} \sin r_\nu q(t-b)]\, dt.$$

Consider the integral

$$\int_a^b F_1(t)\, e^{-r_\nu p(t-b)} \cos r_\nu q(t-b)\, dt.$$

By the second law of the mean it is equal to

$$F_1(b) \int_{\bar{a}}^b e^{-r_\nu p(t-b)} \cos r_\nu q(t-b)\, dt \qquad (a < \bar{a} < b),$$

which after integration becomes

$$\frac{F_1(b)}{r_\nu} [e^{r_\nu p(b-\bar{a})} [q \sin r_\nu q(b-\bar{a}) + p \cos r_\nu q(b-\bar{a})] - p].$$

This expression is less in absolute value than $3F_1(b)/r_\nu$, since p is negative. Similar arguments apply to the three remaining integrals, and to the case where $R(\rho w_i) > 0$.

LEMMA 2. *Let the function* $f(x)$ *have a continuous* m*th derivative of limited variation in the interval* (a, b), *and let* $f(x)$ *and its first* $m - 1$ *derivatives vanish at* a *and at* b. *Then*

I. $R_\nu(x) = O(1/\nu^m)$ *uniformly in* (a, b), *and*

II. $R_\nu(x) = \frac{1}{h} O(1/\nu^{m+1})$ *uniformly in the interval* (a', b'), *where* $a' - a = h$, *and* $b - b' = h$.

First consider the integral

$$\int_a^b f(t)\, u_i(t)\, dt,$$

which by (15) is equal to

$$\int_a^b f(t)\,e^{-\rho w_i(t-c_i')}\,dt + \sum_{j=1}^{m} \frac{1}{(\rho w_i)^j} \int_a^b f(t)\,\psi_j(t)\,e^{-\rho w_i(t-c_i')}\,dt + \frac{1}{\rho^{m+1}} \int_a^b f(t)\,\bar{E}_i\,e^{-\rho w_i(t-c_i')}\,dt. \tag{26}$$

In the first integral of (26) integrate by parts m times. Since $f(t)$ and its first $m-1$ derivatives vanish at both limits of integration, all the integrated terms disappear, giving

$$\int_a^b f(t)\,e^{-\rho w_i(t-c_i')}\,dt = \frac{1}{(\rho w_i)^m} \int_a^b f^{(m)}(t)\,e^{-\rho w_i(t-c_i')}\,dt,$$

where $f^{(m)}(t)$ satisfies the hypotheses of Lemma 1, so that the right hand side is E/r_ν^{m+1}. Similarly after integration by parts $m+1-j$ times we have

$$\frac{1}{(\rho w_i)^j} \int_a^b f(t)\,\psi_j(t)\,e^{-\rho w_i(t-c_i')}\,dt = \frac{1}{(\rho w_i)^{m+1}} \int_a^b \frac{d^{m+1-j}}{dt^{m+1-j}} [f(t)\,\psi_j(t)]\,e^{-\rho w_i(t-c_i')}\,dt.$$

Now $\psi_j(t)$ has a continuous $(m+n-j)$th and therefore at least a continuous $(m+1-j)$th derivative in (a, b), so that

$$\frac{d^{m+1-j}}{dt^{m+1-j}} [f(t)\,\psi_j(t)]$$

must be continuous and therefore bounded in (a, b). Hence if we take absolute values we readily find that the right-hand term of the last equation is of the form E/r_ν^{m+1}. Similar arguments apply to the last term in (26). Therefore

$$\int_a^b f(t)\,u_i(t)\,dt = E/r_\nu^{m+1} \qquad (i = 1, 2, \cdots, n). \tag{27}$$

Now substitute from (27) into (21) and take absolute values. Since $y_i(x)$ is bounded in (a, b) as ρ becomes infinite we have

$$|R_\nu(x)| = \frac{E}{r_\nu^{m+1}} \int_{\gamma_\nu} |d\rho| = E/r_\nu^m = O(1/\nu^m).$$

This proves I.

To prove II write

$$y_j(x) = e^{\rho w_j(x-c_j)}\,E_j,$$

a form that is obviously possible because of (10). Then by (27)

$$|R_\nu(x)| \leqq \sum_{j=1}^{m} \frac{E}{r_\nu^{m+1}} \int_{\gamma_\nu} |e^{\rho w_j(x-c_j)}|\,|w_j\,d\rho|.$$

Let h denote the smaller of the two quantities $|x-a|$ and $|x-b|$. Then

in the notation used previously

$$|R_\nu(x)| \leqq \frac{nE}{r_\nu^m}\left|\int_{\phi_1}^{\phi_2} e^{-r_\nu h \sin\phi}\, d\phi\right|$$
$$\leqq \frac{E'}{r_\nu^m}\left|\int_0^{\pi/2} e^{-2r_\nu h\phi/\pi}\, d\phi\right|$$
$$\leqq \frac{E''}{hr_\nu^{m+1}} = \frac{1}{h}O(1/\nu^{m+1}).$$

LEMMA 3. *If the hypotheses of Lemma 2 are satisfied then*

$$L_\nu(x) = O(1/\nu^{m+1})$$

uniformly in (a, b).

Let

$$F_{jk}(x,t) = \frac{d^k}{dt^k}[f(t)\omega_j(x,t)] \quad (k = 0, 1, \cdots, m-j+1).$$

Then $F_{jk}(x,t)$ vanishes at $t = a$ and at $t = b$ for $k = 0, 1, \cdots, m-j$. Accordingly after integration by parts $(m-j+1)$ times

$$\frac{1}{z^j}\int_a^x F_{j0}(x,t)\, e^{-z(t-x)}\, dt$$
$$= -\sum_{k=0}^{m-j}\frac{F_{jk}(x,x)}{z^{k+j+1}} + \frac{1}{z^{m+1}}\int_a^x F_{j,\,m-j+1}(x,t)\, e^{-z(t-x)}\, dt.$$

Now $F_{j,\,m-j+1}(x,t)$ is of limited variation with respect to t in (a, b) and therefore an easy extension of Lemma 1 shows that the second term on the right in the last equation is E/r_ν^{m+2}. In a similar manner

$$\frac{1}{z^j}\int_x^b F_{j0}(x,t)\, e^{-z(t-x)}\, dt = \sum_{k=0}^{m-j}\frac{F_{jk}(x,x)}{z^{k+j+1}} + E/r_\nu^{m+2}.$$

Therefore

$$-\int_a^x f(t)\,\xi(x,t,z)\, dt = \sum_{j=1}^{m}\sum_{k=0}^{m-j}\frac{F_{jk}(x,x)}{z^{k+j+1}} + E/r_\nu^{m+2},$$
$$\int_x^b f(t)\,\xi(x,t,z)\, dt = \sum_{j=1}^{m}\sum_{k=0}^{m-j}\frac{F_{jk}(x,x)}{z^{k+j+1}} + E/r_\nu^{m+2}.$$

When these expressions are put into (23), and when the two paths of integration are combined into the single circle l, the result is

$$\frac{1}{2\pi\sqrt{-1}}\sum_{j=1}^{m}\sum_{k=0}^{m-j}F_{jk}(x,x)\int_l\frac{dz}{z^{k+j+1}} + E/r_\nu^{m+1} = L_\nu(x).$$

The integral around the closed contour is zero for every pair of values (k, j) in the summation, since $k + j + 1$ is always greater than 1; and therefore

$$L_\nu(x) = E/r_\nu^{m+1} = O(1/\nu^{m+1}).$$

Note. It was solely for the purpose of obtaining this result that it was

found necessary to employ more precise asymptotic forms for the solutions $y_i(x)$ and for the $\bar{y}_i(t)$ than those used by Birkhoff. It was only because of the fact that the functions $\phi_j(x)$ and $\psi_j(x)$ were *independent of* i that $L_\nu(x)$ could be put in the form (23), from which the conclusion of Lemma 3 was derived.

4. Degree of convergence of the series

The results obtained in the preceding sections enable us now to draw certain conclusions concerning the degree of convergence of the series.

Theorem 1. *Let $f(x)$ have a continuous mth derivative of limited variation in (a, b), and let $f(x)$ and its first $m-1$ derivatives vanish at a and at b. Let the coefficients $P_s(x)$ in the differential equation have continuous derivatives of order $(m+n-s)$. Then*

$$S_\nu(x) = f(x) + O(1/\nu^m)$$

uniformly in (a, b).

Under the hypotheses of the theorem we obtain from (25), by Lemma 2 I and Lemma 3,

$$S_\nu(x) = \frac{1}{\pi}\int_a^b f(t)\frac{\sin r_\nu(x-t)}{x-t}dt + O(1/\nu^m),$$

uniformly in (a, b).

Now the particular differential system which yields the Fourier's series is the following:

$$\frac{d^2u}{dx^2} + \lambda u = 0, \tag{28}$$

$$u(a) = u(b), \qquad u'(a) = u'(b),$$

a system that evidently satisfies all the conditions that we have imposed on the system (1) and (2). Moreover, although the characteristic numbers of this system are not simple, it is easily shown by direct computation that the contour integral corresponding to that in (5) formed for the system (28) actually represents the sum of the first ν terms of the Fourier's series. In this case in counting the number of characteristic numbers inclosed in the contour it must be remembered that each one counts double.

From these considerations it follows that if $\Sigma_\nu(x)$ denotes the sum of the first ν terms of the Fourier's series for $f(x)$, then

$$\Sigma_\nu(x) = \frac{1}{\pi}\int_a^b f(t)\frac{\sin \bar{r}_\nu(x-t)}{x-t}dt + O(1/\nu^m).$$

Now when ν is sufficiently large, $\nu > \mu$, we can certainly choose the radii of the paths of integration in the ρ-plane the same for both systems, *i. e.*, make $r_\nu = \bar{r}_\nu$, for $\nu = \mu, \mu+1, \mu+2, \cdots$. For we have

$$\bar{r}_\nu = \frac{\pi\nu}{b-a} + q',$$

and

$$r_\nu = \frac{\pi\nu}{b-a} + q'',$$

where q' is chosen to keep the paths $\bar{\gamma}_\nu$ for the Fourier's series uniformly away from the characteristic numbers of (28), while q'' is chosen to keep the paths γ_ν uniformly away from the numbers ρ_ν. From the distribution of the characteristic numbers for the two systems we infer that when ν is large we may choose $q' = q''$. Then

$$S_\nu(x) - \Sigma_\nu(x) = O(1/\nu^m).$$

But when $f(x)$ satisfies the hypotheses of the theorem it is known that

$$\Sigma_\nu(x) = f(x) + O(1/\nu^m).$$

Therefore

$$S_\nu(x) = f(x) + O(1/\nu^m).$$

THEOREM 2. *Let $f(x)$ have an mth derivative satisfying the Lipschitz condition*

$$|f^{(m)}(x_2) - f^{(m)}(x_1)| \leqq N|x_2 - x_1|$$

in (a, b), and let $f(x)$ and its first $m-1$ derivatives vanish at a and at b. Let the coefficients $P_s(x)$ have continuous derivatives of order $(m+n-s)$ in (a, b). Then

$$S_\nu(x) = f(x) + O\left(\frac{\log \nu}{\nu^{m+1}}\right)$$

uniformly in (a', b'), where $a' - a = b - b' = O(1/\log \nu)$.

Any function satisfying a Lipschitz condition will also be continuous and of limited variation. Therefore Lemmas 2 II and 3 hold for $f(x)$ with the new hypotheses, and consequently

$$L_\nu(x) = O(1/\nu^{m+1}),$$

and

$$R_\nu(x) = \frac{1}{h} O(1/\nu^{m+1}) = O\left(\frac{\log \nu}{\nu^{m+1}}\right),$$

provided that

$$h = O(1/\log \nu).$$

Hence

$$S_\nu(x) - \Sigma_\nu(x) = O\left(\frac{\log \nu}{\nu^{m+1}}\right).$$

But with the given hypotheses it is known* that

$$\Sigma_\nu(x) = f(x) + O\left(\frac{\log \nu}{\nu^{m+1}}\right).$$

* D. Jackson, these Transactions, vol. 13 (1912), pp. 491–515, Theorem V, and vol. 14 (1913), pp. 343–366, Theorem X.

Therefore

$$S_\nu(x) = f(x) + O\left(\frac{\log \nu}{\nu^{m+1}}\right)$$

uniformly in (a', b'), where $a' - a = b - b' = h = O(1/\log \nu)$.

In Theorem 2 nothing is said about the constant implied in the O-notation. A result which is more precise in this respect is as follows:

THEOREM 3. *Let $f(x)$ satisfy the hypotheses of Theorem 2, but let the coefficients $P_s(x)$ have continuous derivatives of order $(m + n - s + 1)$ in (a, b). Then for x in (a', b') and for ν sufficiently large*

$$|S_\nu(x) - f(x)| \leqq C_m N \frac{\log \nu}{\nu^{m+1}},$$

where C_m is a constant depending only on m. Here $a' - a = b - b' = h > 0$, where h is independent of ν.

With the hypotheses of the present theorem the asymptotic formulas for y_i and $\bar{y}_i$ can be carried out to terms in $1/\rho^{m+2}$. It will therefore follow from the reductions of § 2 that

$$S_\nu(x) = \frac{1}{\pi}\int_a^b f(t)\frac{\sin r_\nu(x-t)}{x-t}dt + L_\nu(x) + R_\nu(x) + O\left(\frac{\log \nu}{\nu^{m+2}}\right).$$

From Lemma 2 II and Lemma 3 (both of which hold of course under the present hypotheses), we have

$$S_\nu(x) = \frac{1}{\pi}\int_a^b f(t)\frac{\sin r_\nu(x-t)}{x-t}dt + O(1/\nu^{m+1}),$$

when x is in (a', b') interior to (a, b). Therefore

$$S_\nu(x) - \Sigma_\nu(x) = O(1/\nu^{m+1}).$$

But with the hypotheses of the theorem it is known* that

$$|\Sigma_\nu(x) - f(x)| \leqq \tfrac{1}{2} C_m N \frac{\log \nu}{\nu^{m+1}},$$

where C_m depends only on m. Therefore when ν is large enough

$$|S_\nu(x) - f(x)| \leqq C_m N \frac{\log \nu}{\nu^{m+1}},$$

for x in (a', b').

BOWDOIN COLLEGE,
January, 1917

* See next preceding footnote.

PROBLEMS IN THE THEORY OF ORDINARY LINEAR DIFFERENTIAL EQUATIONS WITH AUXILIARY CONDITIONS AT MORE THAN TWO POINTS*

BY

CHARLES E. WILDER

In this paper is considered a differential system consisting of an ordinary linear differential equation and auxiliary conditions involving linearly the values of the solution and its derivatives at interior points, as well as at the end points, of the interval over which the equation is considered. The notions of Green's function and adjoint systems, already introduced by Birkhoff† and further developed by Bôcher‡ for auxiliary conditions involving the end points only, are extended to the more general system. The problem of the expansion of arbitrary functions in terms of the characteristic solutions of the system is stated, but the convergence proof is given in another paper.§

The author wishes to express his gratitude to Professor Bôcher at whose suggestion he undertook the work, and to Professor Jackson under whose guidance most of it has been done.

1. The auxiliary condition problem

Consider the differential expression

$$L(u) \equiv \frac{d^n u}{dx^n} + p_1(x)\frac{d^{n-1} u}{dx^{n-1}} + \cdots + p_{n-1}(x)\frac{du}{dx} + p_n(x)u$$

and the equation $L(u) = p(x)$, in which $p(x)$, $p_1(x)$, $\cdots$, $p_n(x)$, are functions of the real variable x, continuous together with their derivatives of all orders in the closed interval (a, b). Let $a_1 = a, a_2, a_3, \cdots, a_k = b$, be k points of this interval arranged in order of ascending algebraical magnitude. If $\phi(x)$ is any function of x which has at these points derivatives of the first

* Presented to the Society, April 29, 1916, as part of a more extensive paper, with a slightly different title; the remainder of the paper has already been published in these Transactions, vol. 18 (1917), pp. 415–442.

† Birkhoff, these Transactions, vol. 9 (1908), pp. 373–395.

‡ Bôcher, these Transactions, vol. 14 (1913), pp. 403–420.

§ See first footnote above.

$n-1$ orders, then let

$$W_{ij}(\phi) = \alpha_{ij}\phi(a_j) + \alpha'_{ij}\phi'(a_j) + \alpha''_{ij}\phi''(a_j) + \cdots + \alpha_{ij}^{(n-1)}\phi^{(n-1)}(a_j) \qquad (j = 1, 2, \cdots, k),$$

and let

$$W_i(\phi) = W_{i1}(\phi) + W_{i2}(\phi) + \cdots + W_{ik}(\phi),$$

in which $\alpha_{ij}, \alpha'_{ij}, \alpha''_{ij}, \cdots, \alpha_{ij}^{(n-1)}$, are constants, real or complex.

If π_i represents a real or complex number, then the system

$$L(u) = p, \qquad W_i(u) = \pi_i \qquad (i = 1, 2, 3, \cdots, n),$$

consisting of the differential equation with n *auxiliary conditions*, is called the complete system; the system

$$L(u) = p, \qquad W_i(u) = 0 \qquad (i = 1, 2, 3, \cdots, n), \tag{1}$$

is called the semi-homogeneous system; and finally the system

$$L(u) = 0, \qquad W_i(u) = 0 \qquad (i = 1, 2, 3, \cdots, n), \tag{2}$$

is called the reduced system.

The reduced system is said to be *compatible* if it has solutions other than zero, otherwise it is *incompatible*. It has *r-fold compatibility* if it has r and only r linearly independent solutions.* By direct substitution of the general solution of $L(u) = 0$ in the auxiliary conditions $W_i(u) = 0$ it is found that

If $y_1, y_2, \cdots, y_n$, is a fundamental system of solutions of $L(u) = 0$, a necessary and sufficient condition that the reduced system have r-fold compatibility is that the following determinant be of rank $n - r$:

$$D \equiv |W_i(y_j)| \qquad (i, j = 1, 2, \cdots, n). \tag{3}$$

If we denote by u_0 a particular solution of $L(u) = p$, then it is found in the same way that

If the reduced system has r-fold compatibility, a necessary and sufficient condition that the complete system have a solution is that the matrix obtained by bordering the determinant D on the right by the quantities $W_i(u_0) - \pi_i$, $i = 1, 2, \cdots, n$, be of rank $n - r$.

Whence, in particular, for $r = 0$,

If the reduced system is incompatible, the complete system always has a solution.

That this solution is unique follows at once from the fact that the difference of any two solutions of the complete system is a solution of the reduced system.

Assuming that the auxiliary conditions are linearly independent, we shall now define a function, $G(x, s)$, called the *Green's function* for the reduced system, by requiring that for every value of s in the interval (a, b), it shall have as a function of x the following properties:

* Bôcher, loc. cit., p. 405.

I. It is continuous in the interval (a, b), together with its first $n-2$ derivatives;

II. It satisfies $L(u) = 0$ at all points of the interval except the point s, where its $(n-1)$th derivative has a finite jump of magnitude 1;

III. It satisfies the auxiliary conditions $W_i(G) = 0, i = 1, 2, \cdots, n$.

The proof that these properties completely define the function is as follows:

Since $G(x, s)$ satisfies $L(u) = 0$, it can be written in terms of a fundamental system of solutions,

$$G(x,s) = c_1(s)y_1(x) + c_2(s)y_2(x) + \cdots + c_n(s)y_n(x), \qquad x \leqq s;$$

$$= d_1(s)y_1(x) + d_2(s)y_2(x) + \cdots + d_n(s)y_n(x), \qquad x \geqq s.$$

Now $G(x, s)$ and its first $n-2$ derivatives are continuous at the point s, while the $(n-1)$th derivative has a finite jump of magnitude 1 there, whence

$$c_1(s)y_1(s) + \cdots + c_n(s)y_n(s) = d_1(s)y_1(s) + \cdots + d_n(s)y_n(s),$$

$$c_1(s)y_1'(s) + \cdots + c_n(s)y_n'(s) = d_1(s)y_1'(s) + \cdots + d_n(s)y_n'(s),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,$$

$$c_1(s)y_1^{(n-2)}(s) + \cdots + c_n(s)y_n^{(n-2)}(s) = d_1(s)y_1^{(n-2)}(s) + \cdots + d_n(s)y_n^{(n-2)}(s),$$

$$c_1(s)y_1^{(n-1)}(s) + \cdots + c_n(s)y_n^{(n-1)}(s) = d_1(s)y_1^{(n-1)}(s) + \cdots + d_n(s)y_n^{(n-1)}(s) - 1.$$

If we set $d_i(s) - c_i(s) = z_i(s)$, we can write these equations in the form

$$\sum_{i=1}^{n} z_i(s)y_i^{(j)}(s) = 0 \qquad (j = 0, 1, 2, \cdots, n-2),$$

$$\sum_{i=1}^{n} z_i(s)y_i^{(n-1)}(s) = 1.$$

The determinant of these equations is the wronskian of the y's and so is not zero, since they are linearly independent. Hence the z's are uniquely determined and are continuous functions of s in the interval (a, b). Knowing the z's, we determine the c's and the d's by means of III. If we assume for purposes of proof that s is in the interval (a_l, a_m), we have

$$W_i(G) \equiv \sum_{j=1}^{n}\left(c_j\sum_{q=1}^{l} W_{iq}(y_j) + d_j\sum_{q=m}^{n} W_{iq}(y_j)\right) = 0 \qquad (i = 0, 1, 2, \cdots, n-2),$$

and on substitution of $d_j - z_j$ for c_j this becomes

$$\sum_{j=1}^{n} d_j W_i(y_j) = \sum_{j=1}^{n} z_j\sum_{q=1}^{l} W_{iq}(y_j) \qquad (i = 1, 2, 3, \cdots, n).$$

These are n equations in the n unknowns, the d's, and their determinant is D, so that if D does not vanish the unknowns are uniquely determined, and then by the relation $d_j - z_j = c_j$ the c's are also. Hence if the reduced system is incompatible it possesses one and only one Green's function.

By direct substitution of

$$U(x) = \int_a^b p(s)\, G(x, s)\, ds$$

in the system (1) it is found that the function $U(x)$ is the solution of the semi-homogeneous system, by reason of the three defining conditions for $G(x, s)$.*

If p_i has derivatives of the first $n - i$ orders then the first n derivatives of the functions z_i (called adjoint functions) exist, and satisfy the differential equation

$$M(v) \equiv (-1)^n \frac{d^n v}{ds^n} + (-1)^{n-1} \frac{d^{n-1}(p_1 v)}{ds^{n-1}} + \cdots - \frac{d(p_{n-1} v)}{ds} + p_n v = 0,$$

called the *adjoint equation*.† From this, and the explicit expressions for $G(x, s)$ in terms of the y's, we can deduce easily the following properties of $G(x, s)$ as a function of s:

I. It is continuous in the interval (a, b), together with its first $n - 2$ derivatives, except at the points a_i, where it has finite jumps whose magnitudes are in general functions of x;

II. It satisfies $M(v) = 0$ at all points of the interval (a, b), except the above, and the point x, where its $(n - 1)$th derivative has a finite jump of magnitude $(-1)^{n-1}$.

Since $G(x, s)$ as a function of s has discontinuities it is apparent that an *adjoint system* in the accepted sense of the term‡ does not exist. In considering the expansion problem it becomes necessary to have functions analogous to the solutions of the adjoint system in the two point case. These may be obtained as the solutions of the adjoint to the integral equation which is equivalent to the differential system, but it seems worth while to indicate briefly how they may be found without the aid of the theory of integral equations.

The fact that the Green's function as a function of its second argument has discontinuities at the points a_i suggests that it might be profitable to investigate the case where it has similar discontinuities with regard to its first argument. Such a Green's function would be obtained if the auxiliary

* It is to be noticed that down to this point the operators W_i might have been any linear operators, so that the analysis has a considerably wider range of applicability than has been indicated.

† Schlesinger, *Handbuch der Theorie der linearen Differentialgleichungen*, vol. I, p. 63.

‡ Birkhoff, loc. cit., p. 375, or Bôcher, loc. cit., p. 405.

conditions were such as to require a different solution of the equation between each two points of the set a_i, $i = 1, 2, \cdots, k$. We can define a set of conditions which will do this as follows: Let w_{ij} be of the same form as W_{ij}, but with different values for the α's, and let W_{ij} be considered as applying to the solution to the right of the point a_j, while w_{ij} applies to the solution to the left of that point. Then the auxiliary conditions are

$$\overline{W}_i(u) \equiv W_{i1}(u) + w_{i2}(u) + W_{i2}(u) + \cdots + w_{i,k-1}(u) + W_{i,k-1}(u) + w_{ik}(u) \qquad (i = 1, 2, \cdots, n(k-1)).$$

When the general solution of $L(u) = 0$ is substituted into these auxiliary conditions the constants are determined differently between each two of the points a_i. If that part of the condition W_i which applies to the interval (a_l, a_m), i. e., $W_{il}(u) + w_{im}(u)$, be denoted by $\overline{W}_{il}$, then the determinant whose vanishing is the necessary and sufficient condition that the reduced system have solutions, other than the one which is identically zero throughout the interval (a, b), becomes

$$(4) \quad D \equiv \begin{vmatrix} \overline{W}_{11}(y_1) & \cdots & \overline{W}_{1,k-1}(y_1) \\ \overline{W}_{21}(y_1) & \cdots & \overline{W}_{2,k-1}(y_1) \\ \cdots & \cdots & \cdots \\ \overline{W}_{n(k-1),1}(y_1) & \cdots & \overline{W}_{n(k-1),k-1}(y_1) \end{vmatrix} \begin{vmatrix} \overline{W}_{11}(y_2) & \cdots & \overline{W}_{1,k-1}(y_2) & \cdots \\ \overline{W}_{21}(y_2) & \cdots & \overline{W}_{2,k-1}(y_2) & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \overline{W}_{n(k-1),1}(y_2) & \cdots & \overline{W}_{n(k-1),k-1}(y_2) & \cdots \end{vmatrix} \cdots \begin{vmatrix} \overline{W}_{11}(y_n) & \cdots & \overline{W}_{1,k-1}(y_n) \\ \overline{W}_{21}(y_n) & \cdots & \overline{W}_{2,k-1}(y_n) \\ \cdots & \cdots & \cdots \\ \overline{W}_{n(k-1),1}(y_n) & \cdots & \overline{W}_{n(k-1),k-1}(y_n) \end{vmatrix}.$$

As in the case already considered, one easily finds that the necessary and sufficient condition that the complete system have one and only one solution is that the reduced system be incompatible.

If it is assumed that the auxiliary conditions are linearly independent, the Green's function for the present system is defined for all values of x and s in (a, b) by requiring that for every value of s it shall satisfy as a function of x the following conditions:

I. It is continuous in the interval (a, b), together with its first $n-2$ derivatives, except at the points a_i;

II. Apart from these points, it satisfies $L(u) = 0$ throughout the interval except at the point s, where its $(n-1)$th derivative has a finite jump of magnitude 1;

III. It satisfies the auxiliary conditions $\overline{W}_i(u) = 0$ $(i = 1, 2, \cdots, n(k-1))$.

The proof that these three conditions completely define $G(x, s)$ is parallel to that given for the simpler case.

It is now possible to define adjoint auxiliary conditions, which, taken with the adjoint equation, form the adjoint system. The differential expressions $L(u)$ and $M(v)$ are connected by Lagrange's identity,

$$vL(u) - uM(v) = \frac{d}{dx}P(u, v),$$

where P is a non-singular bilinear form in the $2n$ variables $u, u', u'', \cdots, u^{(n-1)}; v, v', v'', \cdots, v^{(n-1)}$, for every x in the interval. If we integrate this over the interval (a, b), we can for the purposes of integration divide it up into the sum of the integrals taken between each two successive points of discontinuity of G, i. e., the points a_i, and thus admit the possibility of using solutions of the two equations with finite jumps at these points. This gives us

$$\int_a^b [vL(u) - uM(v)]\,dx = P, \tag{5}$$

where P is now a bilinear form in the two sets of $2n(k-1)$ variables each

$$u^{(j)}(a_1^+),\ u^{(j)}(a_2^-),\ u^{(j)}(a_2^+),\ u^{(j)}(a_3^-),\ \cdots,\ u^{(j)}(a_{k-1}^+),\ u^{(j)}(a_k^-),$$

$$v^{(j)}(a_1^+),\ v^{(j)}(a_2^-),\ v^{(j)}(a_2^+),\ v^{(j)}(a_3^-),\ \cdots,\ v^{(j)}(a_{k-1}^+),\ v^{(j)}(a_k^-)$$

$$(j = 0, 1, 2, \cdots, n-1).$$

If $\overline{W}_i(u)$, $i = 1, 2, \cdots, 2n(k-1)$, be any $2n(k-1)$ linearly independent linear forms in the U's, then the U's can be expressed linearly in terms of them, and P becomes a linear form in the $\overline{W}$'s with coefficients which are linear in the v's and which we may denote by $\overline{V}_i$. So we can write

$$P \equiv \overline{W}_1\overline{V}_{2n(k-1)} + \overline{W}_2\overline{V}_{2n(k-1)-1} + \cdots + \overline{W}_{2n(k-1)}\overline{V}_1.$$

In particular we may choose as the first $n(k-1)$ of the linear forms in the u's the left-hand members of our $n(k-1)$ auxiliary conditions. Then the $\overline{V}$'s with indices 1 to $n(k-1)$ when set equal to zero give the adjoint auxiliary conditions. It is easily seen that different choices of the last $n(k-1)$ of the $\overline{W}$'s would lead merely to linear combinations of these adjoint conditions,

so that the system thus defined is unique. From the symmetry in the work it follows that the adjoint relation is a reciprocal one.

By definition $G(x, s)$ as a function of x satisfied the reduced auxiliary conditions. We shall now prove that as a function of s it satisfies the adjoint auxiliary conditions. By substituting, in the integral and in P, $G(x, s)$ for u and $G(s, x)$ for v, we obtain

$$\sum_{j=1}^{n(k-1)} W_{2n(k-1)-j}(G) V_j(G) = 0.$$

The $\overline{W}$'s here appearing are the last $n(k-1)$. We shall make $n(k-1)$ different choices of these. We obtain in this way $n(k-1)$ different sets of $\overline{V}$'s, but as we have already remarked, any set can be expressed as linear combinations of the first set. If we do this we have $n(k-1)$ linear homogeneous equations in the $n(k-1)$ variables, the first set of $\overline{V}$'s. The determinant of these equations depends on the choice of the $\overline{W}$'s, but it is easily shown that these can be chosen in such a way as to make the determinant different from zero, and it follows that the $\overline{V}$'s are all zero. That is, the Green's function as a function of its second argument satisfies the adjoint auxiliary conditions. Combining this with what we already know about $G(x, s)$ as a function of its second argument, we see that $(-1)^n G(x, s)$ is, as a function of s, the Green's function of the adjoint system.

If $y_{ij}(x)$ is defined as equal to $y_i(x)$ in the interval (a_j, a_{j+1}) and as zero elsewhere in the interval (a, b), then the n functions y and the $k-1$ intervals give in all $n(k-1)$ linearly independent solutions of $L(u)$, which it is convenient to call a *fundamental system of discontinuous solutions* with regard to the points a_i. Such a system has properties analogous to the properties of a fundamental system of solutions in the ordinary case. Any solution of $L(u) = 0$ having finite jumps at any of the points a_i, but otherwise continuous in (a, b), can be represented as a linear combination of these. We shall call any function which may have finite jumps at any of the points a_i, but satisfies $L(u) = 0$ at all other points of (a, b), a discontinuous solution (with regard to the points a_i). We have then the

THEOREM. *If the reduced system has h linearly independent discontinuous solutions the adjoint system has also.*

Because of the reciprocal relation between the two systems it is sufficient to show that it has at least h. To prove this let $y(x)$ represent any discontinuous solution of $L(u) = 0$, and form a fundamental system of discontinuous solutions, $z_{ij}(x)$ $(i = 1, 2, \cdots, n; j = 1, 2, \cdots, k-1)$. Substitution of these in the integral and in P leads to the following equations:

$$\sum_{j=1}^{n(k-1)} W_{2n(k-1)-j}(y) V_{j1}(z_i) = 0,$$

$$\sum_{j=1}^{n(k-1)} W_{2n(k-1)-j}(y)\, V_{j2}(z_i) = 0,$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\sum_{j=1}^{n(k-1)} W_{2n(k-1)-j}(y)\, V_{j,\,k-1}(z_i) = 0 \qquad (i = 1, 2, \cdots, n).$$

These may be regarded as $n(k-1)$ homogeneous linear equations to determine the constants $W_{2n(k-1)-j}(y)$, $j = 1, 2, 3, \cdots, n(k-1)$. Their determinant is just the determinant whose rank determines the number of linearly independent discontinuous solutions of the adjoint system, as will be seen by comparing it with (4), which is the corresponding determinant for the reduced system. We wish then to show that the rank of this determinant is at most $r - h$, which we do by showing that the system of equations has at least h independent solutions. To this end let us substitute the h linearly independent solutions of the reduced system in place of the solution $y(x)$ in the W's. This gives h solutions of the system of equations, which can easily be shown to be linearly independent. So the theorem is proved.

Of course these solutions which we have been dealing with may in particular cases be continuous over the whole interval (a, b). We can regard our system (2) as a special case in which the first n auxiliary conditions are the conditions there given, and the rest of the set are the $n(k-2)$ conditions that the solution of the system together with its first $n-1$ derivatives be continuous across each of the $k-2$ points a_i interior to the interval (a, b). We have shown, then, how a system adjoint to the system (2) may be defined, and from now on the discussion will be limited to such a system and its adjoint.

2. Characteristic numbers and solutions

From the system (2) is formed the system

$$L(u) + \lambda u = 0, \qquad W_i(u) = 0 \quad (i = 1, 2, 3, \cdots, n), \tag{6}$$

in which λ is a complex parameter with the whole finite plane for its domain. The adjoint system can be written

$$M(v) + \lambda v = 0, \qquad V_i(v) = 0 \quad (i = 1, 2, 3, \cdots, n(k-1)). \tag{7}$$

The solutions of $L(u) + \lambda u = 0$ are all functions of λ, and so the determinant (3) is also. The values of λ for which this determinant vanishes are the values for which the system above will have solutions other than the identically zero one. Such values are called *characteristic numbers*, and the corresponding solutions are called *characteristic solutions*. The characteristic numbers of the system and its adjoint are, by the theorem just proven, the same; and furthermore the number of linearly independent discontinuous solutions corresponding to any one characteristic number is the same for both systems.

If u and v are solutions of (6) and (7) respectively, belonging to different characteristic numbers, then

$$\int_a^b u(x)\,v(x)\,dx = 0,$$

as will readily be seen by substituting them in the integral (5).

If we assume that each characteristic number has but one corresponding characteristic solution, and try to expand an arbitrary function in terms of these solutions, and if we assume further that the series is such that it can be multiplied through by a continuous function and integrated term by term, then we find that

$$f(x) \equiv \sum_{i=1}^{\infty} c_i\,u_i(x), \tag{8}$$

where

$$c_i \equiv \frac{\int_a^b f(x)\,v_i(x)\,dx}{\int_a^b u_i(x)\,v_i(x)\,dx}.$$

Now the Green's function of the system (6) is, of course, a function of λ, and is defined for all values of λ save the characteristic numbers. The fundamental system $y_1, y_2, y_3, \cdots, y_n$, may be chosen analytic in λ, whence $G(x, s; \lambda)$ is also, except possibly at the characteristic numbers. If $\lambda = \lambda_i$ is a characteristic number to which corresponds but one independent characteristic solution, and for which $G(x, s; \lambda)$ has a pole of the first order, then the residue at the pole is

$$\frac{u_i(x)\,v_i(s)}{\int_a^b u_i(x)\,v_i(x)\,dx},$$

and the integral in the denominator is not zero. The proof given by Birkhoff* applies here.

Under these conditions the terms of the expansion (8) are

$$\int_a^b R_i(x, s)f(s)\,ds, \tag{9}$$

in which $R_i(x, s)$ is the residue corresponding to the ith characteristic number, and the sum of the first m terms is

$$\frac{1}{2\pi i}\int_\Gamma\int_a^b G(x, s; \lambda)f(s)\,ds\,d\lambda, \tag{10}$$

where Γ is a contour in the λ-plane enclosing the first m characteristic numbers.

* Birkhoff, loc. cit., p. 378.

In a previous paper* the author has proved the convergence of the integral (10) when suitable restrictions are placed on $f(x)$ and the auxiliary conditions. Of course in the more general cases the expansion (9) cannot be written in the form (8), that is, the residue cannot be expressed linearly in terms of the characteristic and adjoint functions alone. The new functions which must be introduced have been defined by means of integral equations by Goursat,† but apparently have never been defined for the two point case in terms of the differential system. Such a definition has been formulated by the author, and the accompanying theory forms the subject of another paper.

* Wilder, loc. cit.

† Goursat, Annales de la Faculté des Sciences de Toulouse, series 2, vol. 10 (1908), pp. 5–98.

TRANSFORMATIONS OF APPLICABLE CONJUGATE NETS OF CURVES ON SURFACES*

BY

LUTHER PFAHLER EISENHART

In a previous paper† we developed the theory of transformations T of conjugate systems of curves on a surface into similar systems on other surfaces. In the present paper we are concerned with the application of these results to a particular class of conjugate systems, namely those which are applicable to one or more other systems. Thus if S and $\bar{S}$ are two surfaces upon which the corresponding conjugate system is parametric, we say that the parametric *nets* are *applicable* when corresponding first fundamental coefficients are equal. In order to give this definition analytic form, we suppose that the cartesian coördinates of S and $\bar{S}$ are x, y, z and $\bar{x}, \bar{y}, \bar{z}$ respectively. Since we are interested particularly in parametric nets, we designate them by $N(x)$ and $\bar{N}(\bar{x})$, or merely by N and $\bar{N}$ whenever it is not necessary to specify the coördinates. The first fundamental coefficients are defined by

$$
\begin{aligned}
&E = \Sigma\left(\frac{\partial x}{\partial u}\right)^2, \qquad F = \Sigma\frac{\partial x}{\partial u}\frac{\partial x}{\partial v}, \qquad G = \Sigma\left(\frac{\partial x}{\partial v}\right)^2,\\
&\bar{E} = \Sigma\left(\frac{\partial \bar{x}}{\partial u}\right)^2, \qquad \bar{F} = \Sigma\frac{\partial \bar{x}}{\partial u}\frac{\partial \bar{x}}{\partial v}, \qquad \bar{G} = \Sigma\left(\frac{\partial \bar{x}}{\partial v}\right)^2,
\end{aligned} \tag{1}
$$

the symbol Σ indicating throughout the paper the sum of three terms obtained from the three coördinates. Hence the condition for the applicability of the nets N and $\bar{N}$ is

$$
\begin{aligned}
&\Sigma\left(\frac{\partial x}{\partial u}\right)^2 = \Sigma\left(\frac{\partial \bar{x}}{\partial u}\right)^2, \qquad \Sigma\frac{\partial x}{\partial u}\frac{\partial x}{\partial v} = \Sigma\frac{\partial \bar{x}}{\partial u}\frac{\partial \bar{x}}{\partial v},\\
&\Sigma\left(\frac{\partial x}{\partial v}\right)^2 = \Sigma\left(\frac{\partial \bar{x}}{\partial v}\right)^2.
\end{aligned} \tag{2}
$$

The coördinates of N satisfy an equation of the form

$$
\frac{\partial^2 \theta}{\partial u \partial v} = a\frac{\partial \theta}{\partial u} + b\frac{\partial \theta}{\partial v}, \tag{3}
$$

* Presented to the Society, February 23, 1918.

† These Transactions, vol. 18 (1917), pp. 97-124.

where a and b are functions of E, F, G and their derivatives. Consequently the coördinates of $\bar{N}$ satisfy the same equation. We call it the *point equation of the nets.*

Any net $N'(x')$ whose tangents are parallel to the tangents to $N(x)$ at the point with the same parametric coördinates (u, v) is said to be *parallel* to N. Peterson* showed that if N and $\bar{N}$ are applicable nets, the knowledge of a net parallel to either leads by quadratures to the determination of a net parallel to the other, and the new nets N' and $\bar{N}'$ are applicable. We extend this result by showing that each pair of nets parallel to N and $\bar{N}$ leads by quadratures to the determination of another pair of applicable nets N_1 and $\bar{N}_1$ which are T *transforms* of N and $\bar{N}$ respectively. We recall that in a transformation T of a net N the coördinates of a certain parallel net N' are the direction-parameters of the line joining corresponding points on N and the transform N_1, and the further determination of N_1 requires a solution of equation (3). For the transformations now under discussion the corresponding solution of equation (3) is given by a quadrature after a parallel net is known.

Nets are of three types with regard to applicability. Nets of the first type do not admit any applicable nets. Those of the second type admit one applicable net each, whereas a net of the third type admits an infinity of applicable nets. We say that the latter are *permanent in deformation,* and for the sake of brevity call them *permanent nets.* Every net parallel to a permanent net is a permanent net, and each of the infinity of nets applicable to the one is parallel to one of the infinity applicable to the other by the method of Peterson. Suppose now that we have a permanent net N, two applicable nets $\bar{N}$ and $\bar{\bar{N}}$, and the respective parallel and applicable nets N', $\bar{N}'$, $\bar{\bar{N}}'$. By the process described in the preceding paragraph we obtain two transforms N_1 and N_2 of N, in general distinct, such that the corresponding points of N, N_1, N_2 lie on the same line, whose direction-parameters are the coördinates of N'. At the same time we obtain two transforms of $\bar{N}$ and two of $\bar{\bar{N}}$. As N admits an infinity of applicable nets, this process can be extended with the result that in general N and each of its transforms admit an infinity of transforms. We raise the question whether in any case this infinity of transforms are coincident for each of the nets, so that we obtain a permanent net N_1, whose infinity of applicable nets are the T transforms of the nets applicable to N. We refer to this question as Problem A.

Permanent nets belong to the general class of nets whose tangential coördinates satisfy an equation of Laplace with equal invariants. We have established† transformations of nets of this kind into nets of the same kind. When

* *Ueber Curven und Flachen,* Moskau and Leipzig, 1868, p. 106. Also Stäckel, Mathematische Annalen, vol. 49 (1897), p. 255.

† Rendiconti di Palermo, vol. 39 (1915), pp. 153–176.

in particular N is a permanent net, we find transformations furnishing a solution of Problem A.

In the transformations just referred to we did not consider permanent nets for which the curves in one family are represented on the Gauss sphere by one system of its imaginary generators. Drach* solved the problem of the deformation of nets of this kind. We show how in two ways these nets can be transformed into nets of the same kind as a solution of Problem A.

The third type of permanent nets are those whose two families of curves are represented on the Gauss sphere by the imaginary generators. These curves are the minimal curves on a minimal surface. There are no transformations of nets of this kind into similar nets furnishing a solution of Problem A.

1. General transformations of applicable nets

If $N(x)$ is a net whose cartesian coördinates satisfy equation (3), a parallel net $N'(x')$ is defined by

$$\frac{\partial x'}{\partial u} = h\frac{\partial x}{\partial u}, \qquad \frac{\partial x'}{\partial v} = l\frac{\partial x}{\partial v}, \tag{4}$$

where h and l are a pair of solutions of

$$\frac{\partial h}{\partial v} = (l-h)a, \qquad \frac{\partial l}{\partial u} = (h-l)b. \tag{5}$$

If θ is any solution of (3) and θ' is the solution of the point equation of N' given by the quadrature

$$\frac{\partial \theta'}{\partial u} = h\frac{\partial \theta}{\partial u}, \qquad \frac{\partial \theta'}{\partial v} = l\frac{\partial \theta}{\partial v}, \tag{6}$$

then equations of the form

$$x_1 = x - \frac{\theta}{\theta'}x' \tag{7}$$

define a net N_1 in the relation of a transformation T with N, that is the developables of the congruence of lines joining corresponding points on N and N_1 meet the surfaces on which N and N_1 lie in these nets. Conversely the most general transformation T is defined in this way.

From (7) we have by differentiation

$$\frac{\partial x_1}{\partial u} = \frac{\tau}{\theta'^2}\left(x'\frac{\partial \theta}{\partial u} - \theta'\frac{\partial x}{\partial u}\right), \qquad \frac{\partial x_1}{\partial v} = \frac{\sigma}{\theta'^2}\left(x'\frac{\partial \theta}{\partial v} - \theta'\frac{\partial x}{\partial v}\right), \tag{8}$$

where we have put

$$\tau = h\theta - \theta', \qquad \sigma = l\theta - \theta'. \tag{9}$$

*Annales de Toulouse, ser. 2, vol. 10 (1908), pp. 125–164.

Hence the first fundamental coefficients E_1, F_1, G_1 of N_1 are given by

$$E_1 = \frac{\tau^2}{\theta'^2}\left[\sum x'^2 \cdot \left(\frac{\partial\theta}{\partial u}\right)^2 - 2\sum x' \frac{\partial x}{\partial u} \cdot \theta' \frac{\partial\theta}{\partial u} + \theta'^2 E\right],$$

$$(10)\quad F_1 = \frac{\tau\sigma}{\theta'^2}\left[\sum x'^2 \cdot \frac{\partial\theta}{\partial u}\frac{\partial\theta}{\partial v} - \sum x' \frac{\partial x}{\partial u} \cdot \theta' \frac{\partial\theta}{\partial v} - \sum x' \frac{\partial x}{\partial v} \cdot \theta' \frac{\partial\theta}{\partial u} + \theta'^2 F\right],$$

$$G_1 = \frac{\sigma^2}{\theta'^2}\left[\sum x'^2 \cdot \left(\frac{\partial\theta}{\partial v}\right)^2 - 2\sum x' \frac{\partial x}{\partial v} \cdot \theta' \frac{\partial\theta}{\partial v} + \theta'^2 G\right].$$

Suppose now that N admits an applicable net $\bar{N}(\bar{x})$. Since a and b in (5) are functions of the first fundamental coefficients of N, and consequently of $\bar{N}$, a net $\bar{N}'(\bar{x}')$ parallel to $\bar{N}$ is defined by the quadratures

$$(11)\qquad \frac{\partial\bar{x}'}{\partial u} = h\frac{\partial\bar{x}}{\partial u}, \qquad \frac{\partial\bar{x}'}{\partial v} = l\frac{\partial\bar{x}}{\partial v}.$$

It is readily seen that N' and $\bar{N}'$ are applicable, a result due to Peterson.*

Since θ and θ' are corresponding solutions of the point equations of $\bar{N}$ and $\bar{N}'$, a transform $\bar{N}_1$ of $\bar{N}$ is given by

$$(12)\qquad \bar{x}_1 = \bar{x} - \frac{\theta}{\theta'}\bar{x}'.$$

Ordinarily N_1 and $\bar{N}_1$ are not applicable. We seek the conditions under which they are applicable. From (10) and analogous equations for $\bar{N}_1$ we find that for this to be true it is necessary and sufficient that

$$(\sum x'^2 - \sum \bar{x}'^2)\frac{\partial\theta}{\partial u} = 2\left(\sum x' \frac{\partial x}{\partial u} - \sum \bar{x}' \frac{\partial\bar{x}}{\partial u}\right)\theta',$$

$$(\sum x'^2 - \sum \bar{x}'^2)\frac{\partial\theta}{\partial v} = 2\left(\sum x' \frac{\partial x}{\partial v} - \sum \bar{x}' \frac{\partial\bar{x}}{\partial v}\right)\theta',$$

which in consequence of (6) are equivalent to

$$(13)\qquad \theta' = \sum x'^2 - \sum \bar{x}'^2,$$

to within a negligible constant factor. Koenigs called attention to the fact that when two applicable surfaces are referred to the common conjugate system and consequently have the same point equation, a solution of the latter is given by the difference of the sum of the squares of the coördinates of each surface. Hence θ' as given by (13) is a solution of the point equation of N' and $\bar{N}'$, and the corresponding solution of the point equation of N is given by the quadrature

$$(14)\quad \frac{\partial\theta}{\partial u} = 2\left(\sum x' \frac{\partial x}{\partial u} - \sum \bar{x}' \frac{\partial\bar{x}}{\partial u}\right), \qquad \frac{\partial\theta}{\partial v} = 2\left(\sum x' \frac{\partial x}{\partial v} - \sum \bar{x}' \frac{\partial\bar{x}}{\partial v}\right).$$

* Loc. cit.

Hence we have the

THEOREM. *If N and $\bar{N}$ are applicable nets, each net parallel to N determines a pair of applicable nets N_1 and $\bar{N}_1$, which are the respective transforms of N and $\bar{N}$.*

2. PERMANENT NETS

A permanent net, as defined in the Introduction, is characterized by the property that its curves have the same spherical representation as the asymptotic lines of a surface whose total curvature is of the form

$$K = -\frac{1}{\rho^2}, \qquad \rho = \phi(u) + \psi(v),$$

where ϕ and ψ are functions of u and v alone respectively, the parameters referring to the net.* There are three cases to be considered, according as ρ is constant, involves only one parameter, or both parameters. By a suitable choice of the parameters we can reduce the cases to the canonical forms: $\rho = \text{const.}$, $\rho = v$, $\rho = u + v$. We consider these cases in detail, excluding for the present the possibility of the curves in one or both systems on the sphere being the isotropic generators.

1°. When $\rho = a$, a constant, the coefficients of the linear element of the spherical representation are reducible to the form

$$\mathcal{E} = 1, \qquad \mathcal{F} = \cos 2\omega, \qquad \mathcal{G} = 1, \tag{15}$$

the function ω being a solution of

$$2\frac{\partial^2 \omega}{\partial u \partial v} + \sin 2\omega = 0.\dagger \tag{16}$$

If D, D'' and D_k, D_k'' denote the second fundamental coefficients of N and of an applicable net $\bar{N}$, then for any of the three cases

$$D_k = \tanh \phi \cdot D, \qquad D_k'' = \coth \phi \cdot D'', \tag{17}$$

where ϕ is given by

$$\frac{\partial \phi}{\partial u} = \left\{\begin{matrix}12\\2\end{matrix}\right\}' \tanh \phi, \qquad \frac{\partial \phi}{\partial v} = \left\{\begin{matrix}12\\1\end{matrix}\right\}' \coth \phi, \tag{18}$$

the Christoffel symbols being formed with respect to the linear element of the spherical representation.‡

For the case $\rho = a$ we have

$$D_k = \frac{D}{k}, \qquad D_k'' = kD'', \tag{19}$$

* E., p. 340. A reference of this sort is to the author's *Differential Geometry*.
† E., p. 190.
‡ E., p. 340.

where k is a constant such that when $k = 1$ we have the net N. From the general formulas connecting the fundamental coefficients of a surface, namely*

$$\begin{gathered} \mathcal{E}(EG - F^2) = GD^2, \qquad \mathcal{F}(EG - F^2) = -FDD'', \\ \mathcal{G}(EG - F^2) = ED''^2, \end{gathered} \tag{20}$$

we find for the coefficients of the spherical representation of N the expressions

$$\mathcal{E}_k = \frac{1}{k^2}, \qquad \mathcal{F}_k = \cos 2\omega, \qquad \mathcal{G}_k = k^2. \tag{21}$$

Incidentally we observe that the pseudospherical surfaces with these spherical representations of their asymptotic lines are in the relation of transformations of Lie.†

2°. When $\rho = v$, equations (34), E., p. 340 become

$$\left\{\begin{matrix}12\\1\end{matrix}\right\}' = -\frac{1}{2v}, \qquad \left\{\begin{matrix}12\\2\end{matrix}\right\}' = 0, \tag{22}$$

so that the coefficients of the spherical representation of N are reducible to

$$\mathcal{E} = \frac{1}{v}, \qquad \mathcal{F} = -\frac{\partial \psi}{\partial u}, \qquad \mathcal{G} = \frac{\psi}{v}, \tag{23}$$

where ψ is a solution of a partial differential equation of the third order obtained by equating to unity the expression for the Gaussian curvature of the linear element of the spherical representation.

In this case the integral of (18) is of the form

$$\tanh \phi = \sqrt{1 + kv},$$

where k is a constant. Hence we have

$$D_k = \sqrt{1 + kv}D, \qquad D_k'' = \frac{D''}{\sqrt{1 + kv}}, \tag{24}$$

$$\mathcal{E}_k = \frac{1}{v} + k, \qquad \mathcal{F}_k = -\frac{\partial \psi}{\partial u}, \qquad \mathcal{G}_k = \frac{\psi}{v + kv^2}, \tag{25}$$

$$\rho_k = \frac{v}{1 + kv}, \tag{26}$$

where $-1/\rho_k^2$ is the total curvature of the surface whose asymptotic lines have the spherical representation with the coefficients (25).

3°. When $\rho = u + v$, we have

$$\left\{\begin{matrix}12\\1\end{matrix}\right\}' = \frac{-1}{2(u + v)}, \qquad \left\{\begin{matrix}12\\2\end{matrix}\right\}' = \frac{-1}{2(u + v)},$$

* E., p. 200.

† E., p. 289.

from which it follows that we may take

$$(27) \qquad \mathcal{E} = \frac{\partial \psi}{\partial u}(u+v)^{-1}, \qquad \mathcal{F} = -\frac{\partial^2 \psi}{\partial u \partial v}, \qquad \mathcal{G} = \frac{\partial \psi}{\partial v}(u+v)^{-1},$$

the function ψ being determined by the Gauss equation of the sphere as in the former case.

The integral of the corresponding equations (18) is

$$(28) \qquad \tanh \phi = \sqrt{\frac{1+kv}{1-ku}} \equiv \lambda,$$

so that

$$(29) \qquad D_k = \lambda D, \qquad D_k'' = \frac{1}{\lambda} D'',$$

and

$$(30) \qquad \mathcal{E}_k = \frac{\lambda^2}{u+v}\frac{\partial \psi}{\partial u}, \qquad \mathcal{F}_k = -\frac{\partial^2 \psi}{\partial u \partial v}, \qquad \mathcal{G}_k = \frac{1}{\lambda^2(u+v)}\frac{\partial \psi}{\partial v},$$

with the result

$$(31) \qquad \rho_k = \frac{u+v}{(1-ku)(1+kv)},$$

k being a constant such that when $k = 0$ we have the original net N.

3. Transformations Ω

Let X, Y, Z be the direction-cosines of the normal to a net N and W the fourth tangential coördinate so that

$$Xx + Yy + Zz = W.$$

The net N is said to have equal tangential invariants when the equation which these coördinates satisfy has equal invariants. We suppose that N is such a net, the tangential equation being

$$(32) \qquad \frac{\partial^2 \theta}{\partial u \partial v} + \frac{1}{2}\frac{\partial \log \rho}{\partial v}\frac{\partial \theta}{\partial v} + \frac{1}{2}\frac{\partial \log \rho}{\partial u}\frac{\partial \theta}{\partial v} + \sqrt{\mathcal{E}\mathcal{G}} \cos 2\omega \cdot \theta = 0,$$

and the linear element of the spherical representation

$$(33) \qquad d\sigma^2 = \mathcal{E}du^2 + 2\sqrt{\mathcal{E}\mathcal{G}} \cos 2\omega du dv + \mathcal{G}dv^2.$$

If X', Y', Z' and X'', Y'', Z'' denote the direction-cosines of the bisectors of the angles between the parametric lines on the sphere, we find with the aid of the Gauss equations* for the sphere the following:

* E., p. 160.

$$\frac{\partial X}{\partial u} = \sqrt{\mathcal{E}}\,(\sin\omega X' + \cos\omega X''), \qquad \frac{\partial X}{\partial v} = \sqrt{\mathcal{G}}\,(-\sin\omega X' + \cos\omega X''),$$

$$(34)\quad \frac{\partial X'}{\partial u} = -AX'' - \sqrt{\mathcal{E}}\sin\omega X, \qquad \frac{\partial X'}{\partial v} = BX'' + \sqrt{\mathcal{G}}\sin\omega X,$$

$$\frac{\partial X''}{\partial u} = AX' - \sqrt{\mathcal{E}}\cos\omega X, \qquad \frac{\partial X''}{\partial v} = -BX' - \sqrt{\mathcal{G}}\cos\omega X,$$

where

$$(35)\quad A = \frac{\partial\omega}{\partial u} - \sqrt{\frac{\mathcal{E}}{\mathcal{G}}}\sin 2\omega\frac{\partial\log\sqrt{\rho}}{\partial v}, \qquad B = \frac{\partial\omega}{\partial v} - \sqrt{\frac{\mathcal{G}}{\mathcal{E}}}\sin 2\omega\frac{\partial\log\sqrt{\rho}}{\partial u}.$$

The curves on the sphere represent the asymptotic lines on a surface Σ whose total curvature is $-1/\rho^2$.

From the theory of tangential coördinates we know that if W' is any other solution of equation (32), the functions X, Y, Z, W' are the tangential coördinates of a net N' parallel to N. This net leads to transformations T of N further determined by solutions of the point equation of N. In particular one of these transforms, N_1, is defined in terms of its tangential coördinates X_1, Y_1, Z_1, W_1, which are given by

$$(36)\quad \begin{aligned}\frac{\partial}{\partial u}(\sqrt{\rho\rho_1}\,W'X_1) &= \rho\left(W'\frac{\partial X}{\partial u} - X\frac{\partial W'}{\partial u}\right),\\ \frac{\partial}{\partial v}(\sqrt{\rho\rho_1}\,W'X_1) &= -\rho\left(W'\frac{\partial X}{\partial v} - X\frac{\partial W'}{\partial v}\right),\end{aligned}$$

and similar equations when X_1 and X are replaced by Y_1, Y; Z_1, Z; and W_1, W respectively, the function ρ_1 being a function of u and v such that $-1/\rho_1^2$ is the total curvature of the surface Σ_1 with the same spherical representation of its asymptotic lines as the curves of N_1.* We have said that in this case the nets N and N_1 are in the relation of a transformation Ω.

It is convenient to express X_1 in the form

$$(37)\quad X_1 = \cos\sigma X + \sin\sigma\,(\sin\alpha X' - \cos\alpha X''),$$

where as thus defined σ is the angle between the tangent planes to N and N_1.

Suppose now that N is a permanent net. We have shown† that N_1 also is a permanent net, provided that W', σ, and α satisfy the completely integrable system of equations

$$\frac{\partial\sigma}{\partial u} = \cot\frac{\sigma}{2}\frac{\partial\log\rho}{\partial u}, \qquad \frac{\partial\sigma}{\partial v} = -\tan\frac{\sigma}{2}\frac{\partial\log\rho}{\partial v},$$

* These Transactions, l. c., p. 119.

† Rendiconti di Palermo, loc. cit.

$$\frac{\partial \log W'}{\partial u} = \sqrt{\mathcal{E}} \cot\frac{\sigma}{2}\cos(\alpha+\omega) - \frac{\partial \log \rho}{\partial u}\frac{1}{1-\cos\sigma},$$

$$\frac{\partial \log W'}{\partial v} = -\sqrt{\mathcal{G}} \tan\frac{\sigma}{2}\cos(\alpha-\omega) - \frac{\partial \log \rho}{\partial v}\frac{1}{1+\cos\sigma}, \tag{38}$$

$$\frac{\partial \alpha}{\partial u} = A - \sqrt{\mathcal{E}} \cot\frac{\sigma}{2}\sin(\alpha+\omega),$$

$$\frac{\partial \alpha}{\partial v} = -B + \sqrt{\mathcal{G}} \tan\frac{\sigma}{2}\sin(\alpha-\omega).$$

In this case the parallel net of the transformation is of a special kind, and will be called a *special permanent net.*

4. Solutions of problem A when $\rho = a$

A permanent net N and a parallel net N' admit continuous arrays of deforms $N^{(k)}$ and $N^{(k)'}$ such that for the same value of k these nets are parallel. By means of N' and the function

$$\theta_k' = \sum x'^2 - \sum x_k'^2 \tag{39}$$

we obtain a transform $N_1^{(k)}$ of N, and at the same time by means of θ_k' and $N^{(k)'}$ a transform $(N^{(k)})_1$ of $N^{(k)}$ such that $N_1^{(k)}$ and $(N^{(k)})_1$ are applicable. Ordinarily $N_1^{(k)}$ varies with k. If it does not, it is a permanent net admitting the family of applicable nets $(N^{(k)})_1$. We shall show that if we take any permanent net N and for N' one of the special permanent nets as defined in § 3, the transforms $N_1^{(k)}$ coincide for all values of k and thus we have solutions of Problem A.

We denote by σ_k, α_k, w_k solutions of equations (38) defining a special net $N^{(k)'}$ parallel to $N^{(k)}$. From (15) and (21) we see that ω is independent of k. Hence for $\rho = a$ we have, since $\rho_k = a$ also,

$$\frac{\partial \sigma_k}{\partial u} = 0, \qquad \frac{\partial \sigma_k}{\partial v} = 0,$$

$$\frac{\partial \log w_k}{\partial u} = \frac{1}{k}\cot\frac{\sigma_k}{2}\cos(\alpha_k+\omega),$$

$$\frac{\partial \log w_k}{\partial v} = -\frac{1}{k}\tan\frac{\sigma_k}{2}\cos(\alpha_k-\omega), \tag{40}$$

$$\frac{\partial \alpha_k}{\partial u} - \frac{\partial \omega}{\partial u} + \frac{1}{k}\cot\frac{\sigma_k}{2}\sin(\alpha_k+\omega) = 0,$$

$$\frac{\partial \alpha_k}{\partial v} + \frac{\partial \omega}{\partial v} - k\tan\frac{\sigma_k}{2}\sin(\alpha_k-\omega) = 0.$$

The second fundamental coefficients D_k and D_k'' of the net whose tangential coördinates are X, Y, Z, w_k are given by the general formulas*

$$(41)\qquad \begin{aligned} D_k &= -\left(\frac{\partial^2 w_k}{\partial u^2} - \left\{\begin{matrix}11\\1\end{matrix}\right\}'\frac{\partial w_k}{\partial u} - \left\{\begin{matrix}11\\2\end{matrix}\right\}'\frac{\partial w_k}{\partial v} + \mathcal{E}_k w_k\right), \\ D_k'' &= -\left(\frac{\partial^2 w_k}{\partial v^2} - \left\{\begin{matrix}22\\1\end{matrix}\right\}'\frac{\partial w_k}{\partial u} - \left\{\begin{matrix}22\\2\end{matrix}\right\}'\frac{\partial w_k}{\partial v} + \mathcal{G}_k w_k\right), \end{aligned}$$

where the Christoffel symbols are formed with respect to the linear element of the spherical representation, its coefficients having the values (21). Making use of (40) we find

$$(41')\qquad \begin{aligned} D_k &= \frac{w_k}{k\sin\frac{\sigma_k}{2}}\left[\frac{\cos(\alpha_k-\omega)}{\sin 2\omega}\frac{2\frac{\partial\omega}{\partial u}}{\cos\frac{\sigma_k}{2}} - \frac{1}{k\sin\frac{\sigma_k}{2}}\right], \\ D_k'' &= -\frac{kw_k}{\cos\frac{\sigma_k}{2}}\left[\frac{\cos(\alpha_k+\omega)}{\sin 2\omega}\frac{2\frac{\partial\omega}{\partial v}}{\sin\frac{\sigma_k}{2}} + \frac{k}{\cos\frac{\sigma_k}{2}}\right]. \end{aligned}$$

The values for the net N' are given by (41′) when $k = 1$. It is readily found that the conditions (19) are satisfied, if we take

$$(42)\qquad \cot\frac{\sigma_k}{2} = k\cot\frac{\sigma}{2}, \qquad \alpha_k = \alpha, \qquad w_k = \frac{kw}{\sin^2\frac{\sigma}{2} + k^2\cos^2\frac{\sigma}{2}}.$$

Moreover, these values are consistent with (40).

It remains for us to calculate the expression for θ_k' as defined by (39). The cartesian coördinates x, y, z, of a surface, whose tangential coördinates are X, Y, Z, W, are given by expressions of the form

$$x = WX + \frac{1}{\mathcal{E}\mathcal{G} - \mathcal{F}^2}\left[\mathcal{G}\frac{\partial W}{\partial u}\frac{\partial X}{\partial u} - \mathcal{F}\left(\frac{\partial W}{\partial u}\frac{\partial X}{\partial v} + \frac{\partial W}{\partial v}\frac{\partial X}{\partial u}\right) + \mathcal{E}\frac{\partial W}{\partial v}\frac{\partial X}{\partial v}\right].$$

From this it is readily found that

$$(43)\qquad \sum x^2 = W^2 + \frac{1}{\mathcal{E}\mathcal{G} - \mathcal{F}^2}\left[\mathcal{G}\left(\frac{\partial W}{\partial u}\right)^2 - 2\mathcal{F}\frac{\partial W}{\partial u}\frac{\partial W}{\partial v} + \mathcal{E}\left(\frac{\partial W}{\partial v}\right)^2\right].$$

Applying this to the case where w_k and w are given by (40) and (42), we find

$$\theta_k' = \frac{w^2\sin 2\alpha}{\sin 2\omega}\cdot\frac{\sin^2\frac{\sigma}{2} + k^2\cos\sigma - k^4\cos^2\frac{\sigma}{2}}{\left(\sin^2\frac{\sigma}{2} + k^2\cos^2\frac{\sigma}{2}\right)^2}.$$

* E., p. 164.

Hence θ_k' varies with k only in a constant factor, so that from (6) and (7) it follows that the transform N_1 does not vary with k.

5. Solutions of Problem A when $\rho = v$ and when $\rho = u + v$

We consider next the case when equations (22) and (23) hold. The integral of the first two of equations (38) is

$$\tan^2 \frac{\sigma_k}{2} = \frac{b(1+kv)}{v-b}, \tag{44}$$

where b is a constant, and the other equations become

$$\begin{aligned}
&\frac{\partial \log w_k}{\partial u} = \sqrt{\frac{v-b}{vb}} \cos(\alpha_k + \omega), \\
&\frac{\partial \log w_k}{\partial v} = -\sqrt{\psi}\sqrt{\frac{b}{v^2-vb}} \cos(\alpha_k - \omega) - \frac{bk+1}{2(v-b)(1+kv)}, \\
&\frac{\partial \alpha_k}{\partial u} - \frac{\partial \omega}{\partial u} + \frac{1}{2v\sqrt{\psi}} \sin 2\omega + \sqrt{\frac{v-b}{vb}} \sin(\alpha_k + \omega) = 0, \\
&\frac{\partial \alpha_k}{\partial v} + \frac{\partial \omega}{\partial v} - \sqrt{\psi}\sqrt{\frac{b}{v^2-bv}} \sin(\alpha_k - \omega) = 0.
\end{aligned} \tag{45}$$

From (41) we have in this case

$$-\frac{D_k}{w_k}\frac{1}{1+bk} = \frac{1}{b} + \frac{1 - 2v^2\dfrac{\partial^2\psi}{\partial u^2}}{2v\left[\psi - v^2\left(\dfrac{\partial\psi}{\partial u}\right)^2\right]}\left(\frac{1}{2(v-b)} + \frac{\sqrt{\psi}v\cos(\alpha_k - \omega)}{\sqrt{bv(v-b)}}\right),$$

$$\begin{aligned}
-\frac{D_k''}{w_k}\frac{1}{1+bk} = {}& \frac{3}{2}\cos(\alpha_k - \omega)\sqrt{\frac{b}{v}}\frac{\sqrt{\psi}}{(v-b)^{\frac{3}{2}}} + \frac{1}{4(v-b)^2}\left(4 - \frac{b}{v}\right) \\
&+ \sqrt{\frac{bv}{v-b}}\frac{1}{\left[\psi - v^2\left(\dfrac{\partial\psi}{\partial u}\right)^2\right]}\left(v\psi\frac{\partial^2\psi}{\partial u\partial v} + \psi\frac{\partial\psi}{\partial u} - \frac{v}{2}\frac{\partial\psi}{\partial u}\frac{\partial\psi}{\partial v}\right)\cos(\alpha_k + \omega) \\
&+ \frac{1}{4(v-b)}\frac{1}{\dfrac{\psi}{v^2} - \left(\dfrac{\partial\psi}{\partial u}\right)^2}\left(\frac{1}{v^2}\frac{\partial\psi}{\partial v} - 2\frac{\partial\psi}{\partial u}\frac{\partial^2\psi}{\partial u\partial v} - \frac{2\psi}{v}\right).
\end{aligned}$$

A comparison of these results and (24) shows that these conditions are met, if we treat b as independent of k and take

$$\alpha_k = \alpha, \qquad w_k = \frac{\sqrt{1+kv}}{1+kb}\, w. \tag{46}$$

The net N' corresponds to the value $k = 0$, and w for N' is given by (45) for this value of k. Moreover, the latter are consistent with (45) and (46).

Making use of (43), we find

$$\theta_k' = -\frac{w^2 kv}{\sin 2\omega\,(bk+1)}\left(\sin 2\alpha + \frac{1}{\sqrt{\psi}}\sqrt{\frac{b}{v^2-bv}}\sin(\alpha+\omega)\right). \tag{47}$$

Hence θ_k/θ_k' is independent of k, and all the transforms N_1 coincide.

For the case where the coefficients of the linear element of the spherical representation have the form (30) the equations analogous to (44) and (45) are

$$\tan^2\frac{\sigma_k}{2} = \lambda^2\frac{u+b}{v-b}, \tag{48}$$

and

$$\frac{\partial \log w_k}{\partial u} = \frac{\sqrt{\frac{\partial\psi}{\partial u}}}{\sqrt{u+v}}\sqrt{\frac{v-b}{u+b}}\cos(\alpha_k+\omega) - \frac{bk+1}{2(u+b)(1-ku)},$$

$$\frac{\partial \log w_k}{\partial v} = -\frac{\sqrt{\frac{\partial\psi}{\partial v}}}{\sqrt{u+v}}\sqrt{\frac{u+b}{v-b}}\cos(\alpha_k-\omega) - \frac{bk+1}{2(v-b)(1+kv)},$$

$$\frac{\partial\alpha_k}{\partial u} - \frac{\partial\omega}{\partial u} + \sqrt{\frac{\partial\psi}{\partial u}}\frac{\sin 2\omega}{2(u+v)}\Big/\sqrt{\frac{\partial\psi}{\partial v}} + \sqrt{\frac{\partial\psi}{\partial u}}\frac{\sqrt{v-b}}{\sqrt{(u+v)(u+b)}}\sin(\alpha_k+\omega) = 0, \tag{49}$$

$$\frac{\partial\alpha_k}{\partial v} + \frac{\partial\omega}{\partial v} - \sqrt{\frac{\partial\psi}{\partial v}}\frac{\sin 2\omega}{2(u+v)}\Big/\sqrt{\frac{\partial\psi}{\partial u}} - \sqrt{\frac{\partial\psi}{\partial v}}\frac{\sqrt{u+b}}{\sqrt{(u+v)(v-b)}}\sin(\alpha_k-\omega) = 0.$$

Proceeding as in the former cases with the evaluation of D_k and D_k'', we find

$$\frac{D_k(1-ku)}{w_k(1+bk)} = A, \qquad \frac{D_k''(1+kv)}{w_k(1+bk)} = B,$$

where A and B represent certain complicated expressions not involving k, on the assumption that b is independent of k and that $\alpha_k = \alpha$. It is seen that these values are consistent with (29), if we take

$$w_k = \frac{w}{1+bk}\sqrt{(1+kv)(1-ku)},$$

which expression satisfies (49) on the assumption that $w \equiv w_0$. Now

$$\theta_k' = -\frac{kw^2(u+v)}{\sin 2\omega(1+bk)} \frac{1}{\sqrt{\frac{\partial\psi}{\partial u}\cdot\frac{\partial\psi}{\partial v}}}\left(\sin 2\alpha\sqrt{\frac{\partial\psi}{\partial u}\frac{\partial\psi}{\partial v}}\right.$$
$$-\frac{1}{\sqrt{u+v}}\sqrt{\frac{\partial\psi}{\partial v}}\sqrt{\frac{v-b}{u+b}}\sin(\alpha-\omega)$$
$$\left.+\frac{1}{\sqrt{u+v}}\sqrt{\frac{\partial\psi}{\partial u}}\sqrt{\frac{u+b}{v-b}}\sin(\alpha+\omega)\right).$$

As in the two preceding cases θ_k/θ_k' is independent of k.

In view of these results we have the

THEOREM. *Given a family of applicable permanent nets N_i, of any of the three types, and a family of special permanent nets N_i' parallel to the former; each net N_i is transformable by means of its parallel and the function determined by any other net N_j' into a net N_1 which is independent of the choice of N_j'; the family of transforms thus obtained are deforms of one another.*

6. PERMANENT NETS WITH THE CURVES IN ONE FAMILY REPRESENTED ON THE GAUSS SPHERE BY ISOTROPIC GENERATORS

Drach* determined the equations of all nets of this kind. He showed that the coefficients of the spherical representation of such a net are reducible to the form

$$\mathcal{E} = 0, \qquad \mathcal{F} = \frac{2}{(1+uv)^2}, \qquad \mathcal{G} = \frac{2V'}{Vv(1+uv)} + V_1, \tag{50}$$

where V and V_1 are functions of v alone; also that the tangential coördinates X, Y, Z, W, of such a net are expressible in the form

$$X = \frac{a_1}{1+uv} + b_1, \qquad Y = \frac{a_2}{1+uv} + b_2, \qquad Z = \frac{a_3}{1+uv} + b_3,$$
$$W\sqrt{V} = 2\frac{v\phi(u) + u\psi(v)}{1+uv} - \phi' - \psi', \tag{51}$$

where ϕ and ψ are arbitrary functions of u and v alone and the a's and b's are functions of v alone subject to the conditions:

$$\sum_{i=1}^{3} a_i^2 = 0, \qquad \sum_{i=1}^{3} a_i'^2 = \frac{4}{v^2}, \qquad \sum_{i=1}^{3} a_i''^2 = \frac{4}{v^2}\left(V_1 - \frac{5}{4}\frac{V'^2}{V^2} + \frac{2V'}{vV} + \frac{V''}{V}\right),$$
$$b_i = \left(\frac{V'v}{4V} - \frac{1}{2}\right)a_i + \frac{v}{2}a_i' \quad (i = 1, 2, 3),$$

the primes indicating differentiation with respect to the argument.

* Loc. cit.

If X', Y', Z' denote the direction-cosines of the tangent to the curves $u = \text{const.}$ on the sphere, and X'', Y'', Z'', the direction-cosines of the tangents to the orthogonal trajectories of these curves, we have

$$\frac{\partial X}{\partial u} = \frac{\mathcal{F}}{\sqrt{\mathcal{G}}}(X' + iX''), \qquad \frac{\partial X}{\partial v} = \sqrt{\mathcal{G}}X'. \tag{52}$$

Proceeding as in the determination of the formulas (34), we obtain

$$\begin{aligned} \frac{\partial X'}{\partial u} &= -i\frac{V'}{2V}\frac{\mathcal{F}}{\mathcal{G}}X'' - \frac{\mathcal{F}}{\sqrt{\mathcal{G}}}X, & \frac{\partial X'}{\partial v} &= i\left\{\begin{matrix}22\\1\end{matrix}\right\}'\frac{\mathcal{F}}{\mathcal{G}}X'' - \sqrt{\mathcal{G}}X, \\ \frac{\partial X''}{\partial u} &= i\frac{V'}{2V}\frac{\mathcal{F}}{\mathcal{G}}X' - i\frac{\mathcal{F}}{\sqrt{\mathcal{G}}}X, & \frac{\partial X''}{\partial v} &= -i\left\{\begin{matrix}22\\1\end{matrix}\right\}'\frac{\mathcal{F}}{\mathcal{G}}X', \end{aligned} \tag{53}$$

where

$$\left\{\begin{matrix}22\\1\end{matrix}\right\}' = \frac{1}{2\mathcal{F}}\frac{\partial \mathcal{G}}{\partial v} + \frac{\mathcal{G}}{2\mathcal{F}^2}\frac{\partial \mathcal{G}}{\partial u} - \frac{\mathcal{G}}{\mathcal{F}^2}\frac{\partial \mathcal{F}}{\partial v}. \tag{54}$$

The tangential coördinates satisfy the equation

$$\frac{\partial^2 \theta}{\partial u \partial v} = -\frac{V'}{2V}\frac{\partial \theta}{\partial u} - \mathcal{F}\theta, \tag{55}$$

which evidently has equal invariants. Hence the curves on the sphere represent the asymptotic lines on a surface Σ. We are interested in finding the surfaces Σ_1 such that Σ and Σ_1 are the focal surfaces of a W-congruence. We assume that the direction-cosines of the normal to Σ_1 are given by equations of the form (37), where α and σ are to be determined. We recall* that these direction-cosines must satisfy equations of the form

$$\begin{aligned} \frac{\partial}{\partial u}(\sqrt{V\rho_1}\, wX_1) &= -Vw^2\frac{\partial}{\partial u}\left(\frac{X}{w}\right), \\ \frac{\partial}{\partial v}(\sqrt{V\rho_1}\, wX_1) &= Vw^2\frac{\partial}{\partial v}\left(\frac{X}{w}\right), \end{aligned} \tag{56}$$

where w is a solution of (55) and ρ_1 is the function in the tangential equation for Σ_1 analogous to (32). When the expressions (37) are substituted in (56), and it is noted that similar results must hold also for the Y's and Z's, we are led to the following equations of condition:

* E., p. 419.

$$\frac{\partial\sigma}{\partial u} = \frac{i\mathfrak{F}}{\sqrt{\mathcal{G}}}e^{i\alpha}\left(1+\cos\sigma\sqrt{\frac{V}{\rho_1}}\right)-\sqrt{\frac{V}{\rho_1}}\sin\sigma\frac{\partial\log w}{\partial u},$$

$$\frac{\partial\sigma}{\partial v} = -\sqrt{\mathcal{G}}\sin\alpha\left(1-\cos\sigma\sqrt{\frac{V}{\rho_1}}\right)+\sqrt{\frac{V}{\rho_1}}\sin\sigma\frac{\partial\log w}{\partial v},$$

$$\frac{\partial}{\partial u}\log\sqrt{V\rho_1}+\frac{\partial\log w}{\partial u}\left(1-\cos\sigma\sqrt{\frac{V}{\rho_1}}\right)-i\frac{\mathfrak{F}}{\sqrt{\mathcal{G}}}e^{i\alpha}\sin\sigma\sqrt{\frac{V}{\rho_1}}=0,$$

(57)

$$\frac{\partial}{\partial v}\log\sqrt{V\rho_1}+\frac{\partial\log w}{\partial v}\left(1+\cos\sigma\sqrt{\frac{V}{\rho_1}}\right)-\sqrt{\mathcal{G}}\sin\alpha\sin\sigma\sqrt{\frac{V}{\rho_1}}=0,$$

$$\sin\sigma\left(\frac{\partial\alpha}{\partial u}-i\frac{V'}{2V}\frac{\mathfrak{F}}{\mathcal{G}}\right)+\frac{\mathfrak{F}}{\sqrt{\mathcal{G}}}e^{i\alpha}\left(\cos\sigma+\sqrt{\frac{V}{\rho_1}}\right)=0,$$

$$\sin\sigma\left(\frac{\partial\alpha}{\partial v}+i\left\{\begin{matrix}22\\1\end{matrix}\right\}'\frac{\mathfrak{F}}{\mathcal{G}}\right)+\sqrt{\mathcal{G}}\cos\alpha\left(\cos\sigma-\sqrt{\frac{V}{\rho_1}}\right)=0.$$

By means of these equations the first derivatives of X_1 are reducible to

$$\frac{\partial X_1}{\partial u} = -\left(\cos\sigma\frac{\partial\log w}{\partial u}+i\frac{\mathfrak{F}}{\sqrt{\mathcal{G}}}e^{i\alpha}\sin\sigma\right)\sqrt{\frac{V}{\rho_1}}[\cos\sigma X+(\sin\alpha X'$$
$$-\cos\alpha X'')\sin\sigma]-\sqrt{\frac{V}{\rho_1}}\frac{\mathfrak{F}}{\sqrt{\mathcal{G}}}(X'+iX'')+X\sqrt{\frac{V}{\rho_1}}\frac{\partial\log w}{\partial u},$$

$$\frac{\partial X_1}{\partial v} = \left(\cos\sigma\frac{\partial\log w}{\partial v}-\sqrt{G}\sin\alpha\sin\sigma\right)\sqrt{\frac{V}{\rho_1}}[\cos\sigma X$$
$$+(\sin\alpha X'-\cos\alpha X'')\sin\sigma]+\sqrt{\frac{V}{\rho_1}}\sqrt{\mathcal{G}}X'-X\sqrt{\frac{V}{\rho_1}}\frac{\partial\log w}{\partial v}.$$

From these equations we find the following expressions for the coefficients $\mathcal{E}_1$ of the linear element of the spherical representation of Σ_1:

$$\mathcal{E}_1 = \frac{V}{\rho_1}\left[\left(\frac{\partial\log w}{\partial u}\right)^2-\left(\cos\sigma\frac{\partial\log w}{\partial u}+i\frac{\mathfrak{F}}{\sqrt{\mathcal{G}}}e^{i\alpha}\sin\sigma\right)\right].$$

We are interested particularly in the case where the curves $v = \text{const.}$ on the sphere are minimal, that is when $\mathcal{E}_1 = 0$. From the above we have

$$\frac{\partial\log w}{\partial u} = i\frac{\mathfrak{F}}{\sqrt{\mathcal{G}}}e^{i\alpha}\frac{\sin\sigma}{1-\cos\sigma}. \tag{58}$$

Since the tangential equation is to be of the form (55), ρ_1 is a function of v alone. From (58) and the third of (57) we find that $\rho_1 = V$.* Hence equa-

* It should be remarked that the possible change of signs in (58) consistent with $\mathcal{E}_1 = 0$ is merely equivalent to replacing σ by $\sigma+\pi$, and consequently introduces nothing different from the above.

tions (57) reduce to

$$\frac{\partial \sigma}{\partial u} = 0, \qquad \frac{\partial \sigma}{\partial v} = -\tan\frac{\sigma}{2}\cdot\frac{V'}{V},$$

$$\frac{\partial \log w}{\partial u} = i\frac{\mathfrak{F}}{\sqrt{\mathfrak{G}}}e^{i\alpha}\cot\frac{\sigma}{2}, \tag{59}$$

$$\frac{\partial \log w}{\partial v} = \sqrt{\mathfrak{G}}\sin\alpha\tan\frac{\sigma}{2} - \frac{1}{1+\cos\sigma}\frac{V'}{V},$$

$$\frac{\partial \alpha}{\partial u} - \frac{i}{2}\frac{V'}{V}\frac{\mathfrak{F}}{\mathfrak{G}} + \frac{\mathfrak{F}}{\sqrt{\mathfrak{G}}}e^{i\alpha}\cot\frac{\sigma}{2} = 0,$$

$$\frac{\partial \alpha}{\partial v} + i\left\{\begin{matrix}22\\1\end{matrix}\right\}'\frac{\mathfrak{F}}{\mathfrak{G}} - \sqrt{\mathfrak{G}}\cos\alpha\tan\frac{\sigma}{2} = 0. \tag{60}$$

From these we get

$$\frac{\partial}{\partial u}\log we^{i\alpha} = -\frac{V'}{2V}\frac{\mathfrak{F}}{\mathfrak{G}} = \frac{\partial}{\partial u}\log\sqrt{\mathfrak{G}},$$

$$\frac{\partial}{\partial v}\log we^{i\alpha} = \left\{\begin{matrix}22\\1\end{matrix}\right\}'\frac{\mathfrak{F}}{\mathfrak{G}} + i\sqrt{\mathfrak{G}}\tan\frac{\sigma}{2}e^{-i\alpha}. \tag{61}$$

Moreover, from the first two of (59) we find

$$\tan\frac{\sigma}{2} = \sqrt{\frac{a}{V-a}}, \tag{62}$$

where a is a constant. The function w is of the form (51). Substituting this value in the third of (59), we get

$$\sqrt{V}i\frac{we^{i\alpha}}{\sqrt{\mathfrak{G}}}\cot\frac{\sigma}{2} = (1+uv)v\phi' + \psi - v^2\phi - \frac{\phi''}{2}(1+uv)^2. \tag{63}$$

From (62) and the first of (61) we see that the left-hand member of this equation is necessarily a function of v alone. In order that the expression in the right-hand member be a function of v alone, we must have

$$\phi = bu^2 + 2cu + d,$$

where b, c, and d are constants. Then, if we put $\psi - dv^2 + 2cv - b = \overline{V}$, equation (63) reduces to

$$\sqrt{V}iwe^{i\alpha}\cot\frac{\sigma}{2} = \sqrt{\mathfrak{G}}\overline{V},$$

and we have also

$$\sqrt{V}w = \frac{2u\overline{V}}{1+uv} - \overline{V}'.$$

In order that these values may satisfy the second of equations (61), it is necessary that V be a constant, in which case σ also is constant. From these two equations we have

$$e^{-ia} = \frac{2iu \cot\frac{\sigma}{2}}{(1+uv)\sqrt{V_1}} + V_2, \qquad V_2 = -\frac{\overline{V}'}{\overline{V}} \frac{i}{\sqrt{V_1}} \cot\frac{\sigma}{2}.$$

All of the conditions (59) and (60) will be satisfied, if this value satisfies the last of (61). This gives an equation of Riccati for the determination of V_2, whence ψ is found by a quadrature. Hence we have the

THEOREM. *When the asymptotic lines of a surface Σ have the spherical representation with the coefficients* (50), *the determination of other surfaces Σ_1 of the same kind, such that Σ and Σ_1 are the focal surfaces of a W-congruence, requires the integration of a Riccati equation and quadratures; and this is possible only when the coefficient $\mathcal{G}$ is a function of v alone.*

7. SOLUTIONS OF PROBLEM A FOR PERMANENT NETS FOR WHICH $\mathcal{E} = 0$

We apply the results of the preceding section and obtain a solution of Problem A for the nets whose spherical representation has coefficients of the form

$$\mathcal{E} = 0, \qquad \mathcal{F} = \frac{2}{(1+uv)^2}, \qquad \mathcal{G} = V_1.$$

Since in this case we have

$$\{^{12}_{1}\}' = \{^{12}_{2}\}' = 0,$$

equations (17) assume the form (19), and in place of (21) we have

$$\mathcal{E}_k = 0, \qquad \mathcal{F}_k = \mathcal{F}, \qquad \mathcal{G}_k = \mathcal{G}k^2.$$

Proceeding as in § 4, we find that in this case equations (42) hold also and the function θ_k' is given by

$$\theta_k' = w_k^2 e^{2ia} \frac{k^2-1}{\sin^2\frac{\sigma}{2} + k^2\cos^2\frac{\sigma}{2}}.$$

Hence θ_k/θ_k' is independent of k, and we have a solution of Problem A.

Drach* solved the problem of the deformation of permanent nets with the coefficients (50). We recall his results. If R is a general solution of the Riccati equation

$$R' + \tfrac{1}{2}R^2 + \tfrac{1}{2}\Omega = 0,$$

where

$$\Omega = \frac{TV}{1+kV} + \frac{3}{4}\frac{V'^2}{(1+kV)^2 V^2},$$

* Loc. cit.

k being a parameter and R and T functions of v alone, and if we put

$$\rho = e^{\int R dv}, \qquad \sigma = \frac{1}{2}\int \frac{dv}{\rho},$$

$$A_1 = \rho(1-\sigma^2), \qquad A_2 = i\rho(1+\sigma^2), \qquad A_3 = 2\rho\sigma,$$

then the functions X, Y, Z, given by the expressions

$$-2\frac{A_i u}{1+uv} + A_i' - A_i \frac{\partial}{\partial v}\log\sqrt{\frac{1}{V}+k},$$

as i takes the values 1, 2, 3 respectively, are the direction-cosines and W_k, given by

$$W_k = \sqrt{\frac{1}{V}+k}\left(2\frac{v\phi(u)+u\psi(v)}{1+uv} - \phi'(u) - \psi'(v)\right), \tag{64}$$

the fourth tangential coördinate of a net with the coefficients (50). Moreover if in (64) we put

$$\psi = \Psi V/\sqrt{1+kV},$$

where Ψ is any solution of the equation

$$\Psi''' + \Omega\Psi' + \tfrac{1}{2}\Omega'\Psi = \frac{\Phi V}{\sqrt{1+kV}},$$

Φ being an arbitrary function of v, then all the surfaces obtained by varying k are applicable.

The coefficients of the spherical representation of these surfaces are of the form

$$\mathcal{E}_k = 0, \qquad \mathcal{F}_k = \frac{2}{(1+uv)^2},$$

$$\mathcal{G}_k = \frac{V}{1+kV}\left[\frac{2u}{1+uv}\left(\frac{1}{V}\right)' + \left(\frac{1}{V}\right)'' + T\right].$$

We consider in particular the surfaces for which $\psi = 0$. Making use of the general formula (43) and denoting by x_k', y_k', z_k' the cartesian coördinates of the surface with these tangential coördinates, we find

$$\theta_k' = \sum x_k'^2 - \sum x_0'^2 = k\left[\frac{2\phi}{(1+uv)^2}\left((1+uv)v\phi' - v^2\varphi - \tfrac{1}{2}(1+uv)^2\phi''\right) + \left(2\frac{v\phi}{1+uv} - \phi'\right)^2\right].$$

From the form of this expression it follows that the ratio θ_k'/θ_k is independent of k, and consequently by means of these functions θ_k and θ_k' we get a transformation of any of the surfaces with the tangential coördinate (64), which is a solution of Problem A.

9. When the curves in both systems are represented on the sphere by generators

There remains for consideration the case when the curves in both systems of a net are represented on the sphere by imaginary generators, that is when the surface on which the net lies is minimal. In this case we say that the net N is *minimal*. The coördinates of N and $\bar{N}$ are necessarily of the form*

$$x = U + V, \qquad \bar{x} = e^{ik}\,U + e^{-ik}\,V,$$

where k is a constant, and the three functions U of u alone and the three functions V of v alone are subject to the conditions

$$\sum U'^2 = 0, \qquad \sum V'^2 = 0.$$

Parallel nets N' and $\bar{N}'$ are necessarily given by

$$x' = hU + lV, \qquad \bar{x}' = e^{ik}\,hU + e^{-ik}\,lV,$$

where h and l are any unequal constants. Now

$$\theta'_k = h^2(1 - e^{2ik})\sum U^2 + l^2(1 - e^{-2ik})\sum V^2,$$
$$\theta_k = h(1 - e^{2ik})\sum U^2 + l(1 - e^{-2ik})\sum V^2.$$

In order that θ'/θ be independent of k, we must have either $\sum U^2 = 0$, or $\sum V^2 = 0$. From the Weierstrass form† of the functions U and V it follows that only when the U's are constant can they satisfy simultaneously the conditions $\sum U^2 = 0$, $\sum U'^2 = 0$. Hence in all transformations of this sort N_1 varies with k. Making use of formulas (10), we find that in this case the nets N_1 are not minimal.

Princeton University

* E., p. 327.
† E., p. 262.

A NEW INTEGRAL TEST FOR THE CONVERGENCE AND DIVERGENCE OF INFINITE SERIES*

BY

RAYMOND W. BRINK

INTRODUCTION†

A new sequence of integral tests for the convergence and divergence of infinite series has been developed by the author. Some of the tests of this sequence, and the principle by which they may be discovered will be set forth by him in another article. In the present paper it is his desire to give a central one of these tests, together with some of its applications. This particular integral test appears to play the same rôle when the ratio of successive terms is explicitly known, that the Maclaurin-Cauchy test plays when the individual term is explicitly known. In testing a series of the form

$$u_0 + u_1 + u_2 + \cdots,$$

du Bois-Reymond‡ called those tests that make use of the test ratio

$$r_n = u_{n+1}/u_n,$$

tests of the second kind to distinguish them from tests using the general term of the series u_n itself, which he called *tests of the first kind.* Similarly, the integral tests developed in this paper, which involve a function $r(x)$ such that $r(n) = r_n$, may be called *integral tests of the second kind;* while the Maclaurin-Cauchy integral test, involving a function $u(x)$ for which $u(n) = u_n$, is an *integral test of the first kind.*

Integral tests of the second kind thus apply to series for which a function is known that for successive integral values of the variable takes on the successive values of the ratio of one term to the preceding term. Such a series can be written in the following *normal form*:

$$c + ca_0 + ca_0 a_1 + ca_0 a_1 a_2 + \cdots,$$

* Presented to the Society, April, 1916.

† The thanks of the writer are due to Professor Birkhoff for many suggestions furnished by him during the preparation of this paper.

‡ Journal für Mathematik, vol. 76, p. 61.

where $a_n = a(n)$, $a(x)$ being a known function; and it is to such series that the theorems of this paper apply.

The general method employed to discover the tests of the sequence suggests the test set forth in the first section of the present paper. This test may be considered as the fundamental integral test of the second kind. From it are derived the test given in the second section, and the very simple and useful integral test of the third section. The fourth section gives a test applicable to a series of products of rather general form. The fifth extends the fundamental test to multiple series.

1. Fundamental integral test of the second kind

Theorem I. *Given the series*

$$u_0 + u_1 + u_2 + \cdots \qquad (u_n > 0,\ n \geqq \mu).$$

Let $r_n = u_{n+1}/u_n$, *and suppose that from a certain point* $x = \mu$ *on,* $r(x)$ *is a continuous, positive function such that* $r(n) = r_n$, *and suppose that a constant* m *exists, positive or zero, such that* $r(x') \geqq r(x)$ *when* $x' \geqq x + m$. *Then a necessary and sufficient condition for the convergence of the given series is the convergence of the integral*

$$\int_\mu^\infty e^{\int_\mu^x \log r(x)dx}\, dx.$$

Proof. Under the conditions, from a certain point on either $r(x) > 1$ or $r(x) \leqq 1$. Suppose that $r(x) \leqq 1$, $\mu < x$. We take μ to be an integer. Then

$$(1) \qquad \log r_{n-m} \leqq \int_n^{n+1} \log r(x)\, dx \leqq \log r_{n+1+m}, \qquad \mu \leqq n.$$

We write

$$(2) \qquad \int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\, dx = \int_n^{n+1} e^{\overbrace{\int_\mu^{\mu+1} + \int_{\mu+1}^{\mu+2} + \cdots + \int_{n-1}^{n} + \int_n^x} \log r(x)dx}\, dx,$$

where, as elsewhere in this paper, the bar over the signs of integration indicates that the integrals under it have the same integrand. Therefore, by (1),

$$(3) \qquad \int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\, dx \leqq \int_n^{n+1} e^{\log r_{\mu+1+m} + \log r_{\mu+2+m} + \cdots + \log r_{n+m}}\, dx$$
$$= r_{\mu+1+m} \cdot r_{\mu+2+m} \cdot \cdots \cdot r_{n+m} = \frac{1}{u_{\mu+1+m}}\, u_{n+1+m}.$$

Likewise

$$(4) \qquad \int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\, dx \geqq r_{\mu-m} \cdot r_{\mu+1-m} \cdot \cdots \cdot r_{n-m} = \frac{1}{u_{\mu-m}} \cdot u_{n+1-m}.$$

Similar inequalities hold if $r(x) > 1$, $\mu_1 < x$. Since the integral

$$\int_{\mu}^{x} e^{\int_{\mu}^{x} \log r(x)dx}\, dx$$

cannot oscillate, the theorem follows at once from a comparison of the two series

$$\sum_{n=\mu}^{\infty} u_n \qquad \text{and} \qquad \sum_{n=\mu}^{\infty} \int_{n}^{n+1} e^{\int_{\mu}^{x} \log r(x)dx}\, dx$$

by means of (3) and (4).

Example. Discuss the convergence of the series

$$\sum_{n=\mu}^{\infty} u_n,$$

where $u_n = e^{\rho_\mu + \rho_{\mu+1} + \cdots + \rho_{n-1}}$,

$$\rho_n = -\frac{l_1 n \cdot l_2 n \cdot \cdots \cdot l_k n + l_2 n \cdot l_3 n \cdot \cdots \cdot l_k n + \cdots + l_k n + p}{n \cdot l_1 n \cdot l_2 n \cdot \cdots \cdot l_k n},$$

and $l_k x = \log (l_{k-1} x)$, $l_1 x = \log x$. Of course μ is to be taken large enough for all terms of the series to be defined.

For this series $r_n = e^{\rho_n}$, and

$$r(x) = e^{\rho(x)} = e^{-\frac{l_1 x \cdot l_2 x \cdot \cdots \cdot l_k x + l_2 x \cdot l_3 x \cdot \cdots \cdot l_k x + \cdots + l_k x + p}{x \cdot l_1 x \cdot l_2 x \cdot \cdots \cdot l_k x}}.$$

This function satisfies the conditions of the theorem, and we have

$$\int_{x_0}^{x} \log r(x)\, dx = \int_{x_0}^{x} -\frac{l_1 x \cdot \cdots \cdot l_k x + \cdots + l_k x + p}{x l_1 x l_2 x \cdot \cdots \cdot l_k x}\, dx$$

$$= c - \log [x l_1 x \cdots l_{k-1} x\, (l_k x)^p].$$

Therefore

$$\int_{x_0}^{x} e^{\int_{x_0}^{x} \log r(x)dx}\, dx = C \int_{x_0}^{x} \frac{dx}{x \cdot l_1 x \cdot l_2 x \cdot \cdots \cdot l_{k-1} x\, (l_k x)^p}.$$

Hence the given series converges or diverges according as $p > 1$ or $p \leqq 1$.

The theorem can be extended as follows:

THEOREM II. *Given the series*

$$u_0 + u_1 + u_2 + \cdots \qquad (u_n > 0,\ n \geqq \mu).$$

Let $r(x)$ be a function with a continuous derivative $r'(x)$, $x \geqq \mu$, such that $r(n) = r_n = u_{n+1}/u_n$. Suppose that $0 < A \leqq r(x) \leqq B$, and that the integral

$$\int_{\mu}^{\infty} |r'(x)|\, dx$$

converges. Then a necessary and sufficient condition for the convergence of the given series is the convergence of the integral

$$\int_{\mu}^{\infty} e^{\int_{\mu}^{x} \log r(x)dx}\, dx .$$

Proof. Write

$$d_n = \int_n^{n+1} \log r(x)\, dx - \log r_{n+1} \tag{1}$$
$$= \int_n^{n+1} [\log r(x) - \log r_{n+1}] \frac{d}{dx}(x-n)\, dx .$$

Integrate by parts. Then we find

$$d_n = -\int_n^{n+1} (x-n) \frac{r'(x)}{r(x)}\, dx .$$

Therefore

$$|d_n| \leqq \int_n^{n+1} \left| \frac{r'(x)}{r(x)} \right| dx . \tag{2}$$

Now by hypothesis the integral

$$\int_{\mu}^{\infty} |r'(x)|\, dx$$

converges, and therefore the integral

$$\int_{\mu}^{\infty} \left| \frac{r'(x)}{r(x)} \right| dx$$

converges since

$$\left| \frac{r'(x)}{r(x)} \right| \leqq \frac{|r'(x)|}{A} .$$

Consequently the series

$$\sum_{n=\mu}^{\infty} d_n$$

converges absolutely, and we can write

$$\sum_{n=\mu}^{\infty} |d_n| = D .$$

Now

$$\int_n^{n+1} e^{\int_{\mu}^{x} \log r(x)dx}\, dx = \int_n^{n+1} e^{\overbrace{\int_{\mu}^{\mu+1} + \int_{\mu+1}^{\mu+2} + \cdots + \int_{n-1}^{n} + \int_{n}^{x}} \log r(x)dx}\, dx$$
$$= \int_n^{n+1} e^{d_{\mu} + d_{\mu+1} + \cdots + d_{n-1} + \log (r_{\mu+1} \cdot r_{\mu+2} \cdot \cdots \cdot r_n) + \int_n^x \log r(x)dx}\, dx .$$

Since $0 < A \leqq r(x) \leqq B$, we can set $|\log r(x)| \leqq C$. Consequently we

have

$$\frac{e^{-C-D}}{u_{\mu+1}} u_{n+1} \leqq \int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\, dx \leqq \frac{e^{C+D}}{u_{\mu+1}} u_{n+1}.$$

The theorem then follows as before from a comparison of the two series

$$\sum_{n=\mu}^{\infty} u_n \quad \text{and} \quad \sum_{n=\mu}^{\infty} \int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\, dx.$$

The method used in the following *alternative proof* is sometimes useful. Define two functions $r^{(1)}(x)$ and $r^{(2)}(x)$ such that

$$r^{(1)}(x) \cdot r^{(2)}(x) = r(x),$$

and such that $r^{(1)}(x)$ monotonically increases and $r^{(2)}(x)$ monotonically decreases when x increases. We can do this in the following way. Denote $r^{(1)}(n)$ by $r_n^{(1)}$, and $r^{(2)}(n)$ by $r_n^{(2)}$. Take $\log r_\mu^{(1)} = \log r_\mu$ and $\log r_\mu^{(2)} = 0$; if $r'(x) \geqq 0$ take

$$\frac{d \log r^{(1)}(x)}{dx} = \frac{d \log r(x)}{dx}, \qquad \frac{d \log r^{(2)}(x)}{dx} = 0;$$

if $r'(x) \leqq 0$, take

$$\frac{d \log r^{(2)}(x)}{dx} = \frac{d \log r(x)}{dx}, \qquad \frac{d \log r^{(1)}(x)}{dx} = 0.$$

Then

$$\log r^{(1)}(x) + |\log r^{(2)}(x)| = \int_\mu^x \left| \frac{d \log r(x)}{dx} \right| dx + \log r_\mu.$$

Since the integral

$$\int_\mu^\infty \left| \frac{d \log r(x)}{dx} \right| dx = \int_\mu^\infty \left| \frac{r'(x)}{r(x)} \right| dx$$

converges, $\log r^{(1)}(x)$ and $\log r^{(2)}(x)$ are finite, and we can write

$$0 < a < r^{(1)}(x) < b, \qquad 0 < a < r^{(2)}(x) < b.$$

Now

$$\int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\, dx$$

$$= \int_n^{n+1} e^{\int_0^1 [\log r(\mu+x) + \log r(\mu+x+1) + \cdots + \log r(x+n-1)]dx + \int_n^x \log r(x)dx}\, dx$$

$$= \int_n^{n+1} e^{\log [r^{(1)}(\mu+\theta) \cdot r^{(2)}(\mu+\theta) \cdot r^{(1)}(\mu+\theta+1) \cdot r^{(2)}(\mu+\theta+1) \cdots r^{(1)}(n-1+\theta) r^{(2)}(n-1+\theta)] + \int_n^x \log r(x)dx}\, dx,$$

$0 < \theta < 1$. Therefore

$$\int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\,dx \leqq e^C \cdot r_{\mu+1}^{(1)} \cdot r_\mu^{(2)} \cdot r_{\mu+2}^{(1)} \cdot r_{\mu+1}^{(2)} \cdot \cdots \cdot r_n^{(1)} \cdot r_{n-1}^{(2)}$$

$$= e^C \cdot r_\mu^{(2)} \cdot r_n^{(1)} \frac{u_n}{u_{\mu+1}} \leqq \frac{e^C b^2}{u_{\mu+1}} u_n.$$

Likewise

$$\int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\,dx \geqq \frac{e^{-C} a^2}{u_{\mu+1}} u_n.$$

The theorem then follows as before.

Instead of testing for the convergence of the integral

$$\int_\mu^\infty |r'(x)|\,dx$$

it is sometimes convenient to test directly for the convergence of the integral

$$\int_\mu^\infty \left|\frac{r'(x)}{r(x)}\right| dx,$$

which is equivalent.

It is clear that Theorem II is included in the following theorem. We thus have still a third proof of Theorem II.

THEOREM III. *Given the series*

$$u_0 + u_1 + u_2 + \cdots, \qquad (u_n > 0,\ n \geqq \mu).$$

Let $r(x)$ *be a continuous function such that, for* $x \geqq \mu$,

(a) $$r(n) = r_n = \frac{u_{n+1}}{u_n},$$

(b) $$0 < A \leqq r(x) \leqq B,$$

(c) $$|r(x') - r(x)| \leqq f(x),$$

whenever $0 \leqq (x' - x) \leqq 1$, *the series* $\sum_{n=\mu}^{n=\infty} f(n)$ *being a convergent series. Then a necessary and sufficient condition for the convergence of the given series is the convergence of the integral*

$$\int_\mu^\infty e^{\int_\mu^x \log r(x)dx}\,dx.$$

Proof. By condition (c),

(1) $$r(x') \leqq r(x) + f(x) \leqq r(x)\left[1 + \frac{f(x)}{A}\right],$$

$$0 \leqq x' - x \leqq 1, \qquad x \geqq \mu.$$

$$(2) \qquad r(x') \geqq r(x)\left[1 - \frac{f(x)}{A}\right], \qquad 0 \leqq x' - x \leqq 1, \qquad x \geqq \mu.$$

We have

$$\int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\,dx = \int_n^{n+1} e^{\log (r(\mu+\theta)\cdot r(\mu+1+\theta)\cdots r(n-1+\theta)) + \int_n^x \log r(x)dx}\,dx,$$

where $0 < \theta < 1$. As before, we write $|\log r(x)| \leqq \log C$. Then by (1),

$$\int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\,dx \leqq C\left[r_\mu\left(1 + \frac{f(\mu)}{A}\right) r_{\mu+1}\left(1 + \frac{f(\mu+1)}{A}\right) \cdots\cdots r_{n-1}\left(1 + \frac{f(n-1)}{A}\right)\right]$$

$$= \frac{C}{u_\mu} u_n \prod_{n=\mu}^{n-1}\left(1 + \frac{f(n)}{A}\right).$$

Since the series

$$\sum_{n=\mu}^{\infty} f(n)$$

converges, and $f(n) \geqq 0$, the product

$$\prod_{n=\mu}^{\infty}\left(1 + \frac{f(n)}{A}\right)$$

converges to a value F. Therefore

$$\int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\,dx \leqq \frac{CF}{u_\mu} u_n.$$

Similarly, by (2), the product

$$\prod_{n=\mu}^{\infty}\left(1 - \frac{f(n)}{A}\right),$$

converges, say to F_1, and we get

$$\int_n^{n+1} e^{\int_\mu^x \log r(x)dx}\,dx \geqq \frac{F_1}{Cu_\mu} u_n.$$

The theorem then follows by comparison.

2. A DERIVED TEST

The following test is based upon Theorems I, II, and III, in much the same way that Ermakoff's test* is based upon the familiar Maclaurin-Cauchy integral test.

*Bulletin des sciences mathématiques vol. 2 (1871), p. 250.

THEOREM IV. *Given the series*

$$u_0 + u_1 + u_2 + \cdots \qquad (u_n > 0,\ n \geqq \mu).$$

Let $r_n = u_{n+1}/u_n$, *and suppose that* $r(x)$ *is a function satisfying the preliminary conditions given in any one of the Theorems I, II, or III, and that* $r(x) \leqq 1$. *Then the series converges if*

$$\frac{[r(e^x)]^{e^x}}{r(x)} < \nu < \frac{1}{e}, \qquad m \leqq x,$$

and diverges if

$$\frac{[r(e^x)]^{e^x}}{r(x)} > \nu > \frac{1}{e}, \qquad m \leqq x.$$

Proof. Since the preliminary conditions hold for one of the earlier tests, the series converges if and only if the integral

$$\int_\mu^\infty e^{\int_\mu^x \log r(x)dx}\, dx$$

converges. Suppose first that

$$\frac{[r(e^x)]^{e^x}}{r(x)} < \nu < \frac{1}{e}, \qquad \mu \leqq m \leqq x.$$

Then

$$[r(e^x)]^{e^x} < \nu \cdot r(x)$$

and

$$e^x \log r(e^x) < \log \nu + \log r(x).$$

It follows that

$$\int_m^x e^x \log r(e^x)\, dx < (x - m) \log \nu + \int_m^x \log r(x)\, dx, \qquad m < x;$$

or, after a change in the variable of integration,

$$\int_{e^m}^{e^x} \log r(x)\, dx < (x - m) \log \nu + \int_m^x \log r(x)\, dx.$$

Consequently

$$\int_{x_0}^x e^x \cdot e^{\int_{e^m}^{e^x} \log r(x)dx} dx < \frac{1}{\nu^m} \int_{x_0}^x e^x \cdot \nu^x\, e^{\int_m^x \log r(x)dx} dx, \qquad m < x_0 < x.$$

Then, once more changing the variable of integration, we have

$$\int_{e^{x_0}}^{e^x} e^{\int_{e^m}^x \log r(x)dx} dx < \frac{1}{\nu^m} \int_{x_0}^x \frac{1}{\lambda^x} e^{\int_m^x \log r(x)dx} dx, \qquad \lambda > 1.$$

Therefore, since $\log r(x) \leqq 0$,

$$\int_{e^{x_0}}^{e^x} e^{\int_{e^m}^{x} \log r(x)dx}\, dx < \frac{1}{\nu^m}\int_{x_0}^{x} \frac{dx}{\lambda^x}, \qquad m < x_0 < x.$$

But the integral

$$\int_{x_0}^{\infty} \frac{dx}{\lambda^x}$$

converges. Consequently the integral

$$\int_{e^{x_0}}^{\infty} e^{\int_{e^m}^{x} \log r(x)dx}\, dx,$$

and therefore the given series, converge.

Now suppose that

$$\frac{[r(e^x)]^{e^x}}{r(x)} > \nu > \frac{1}{e}, \qquad \mu \leqq m < x.$$

We can take $1/e < \nu < 1$. Then

$$[r(e^x)]^{e^x} > \nu \cdot r(x), \qquad \frac{1}{e} < \nu < 1, \qquad m < x.$$

As before, reversing inequalities, we get

$$\int_{e^{x_0}}^{e^x} e^{\int_{e^m}^{x} \log r(x)dx}\, dx > \frac{1}{\nu^m}\int_{e^{x_0}}^{e^x} \lambda^x\, e^{\int_{m}^{x} \log r(x)dx}\, dx, \qquad \lambda > 1.$$

That is

$$\int_{x}^{e^x} e^{\int_{e^m}^{x} \log r(x)dx}\, dx > \frac{1}{\nu^m}\int_{x_0}^{x} \lambda^x\, e^{\int_{m}^{x} \log r(x)dx}\, dx - \int_{e^{x_0}}^{e^x} e^{\int_{e^m}^{x} \log r(x)dx}\, dx$$

$$> \frac{1}{\nu^m}\int_{x_0}^{x} \lambda^x\, e^{\int_{m}^{x} \log r(x)dx}\, dx - \int_{e^{x_0}}^{e^x} \lambda^x\, e^{\int_{m}^{x} \log r(x)dx}\, dx,$$

$$m' < x_0 < x.$$

Then, since $\nu < 1$,

$$\int_{x}^{e^x} e^{\int_{e^m}^{x} \log r(x)dx}\, dx > \int_{x_0}^{e^{x_0}} \lambda^x\, e^{\int_{m}^{x} \log r(x)dx}\, dx = c > 0.$$

This inequality holds for all values of x beyond a certain point. If we take x_1 some number beyond this point, and set

$$x_2 = e^{x_1}, \qquad x_3 = e^{x_2}, \qquad x_4 = e^{x_3}, \qquad \cdots,$$

each term of the series

$$\int_{x_1}^{x_2} e^{\int_{c_m}^{x} \log r(x)dx}\, dx + \int_{x_2}^{x_3} e^{\int_{c_m}^{x} \log r(x)dx}\, dx + \cdots$$

is greater than the positive constant c, and this series therefore diverges. Therefore the integral

$$\int_{x_0}^{\infty} e^{\int_{c_m}^{x} \log r(x)dx}\, dx,$$

and consequently the given series, diverges.

By the transformation $x = e^{x'}$, the test is changed to the following form, sometimes more convenient: *The series is convergent if*

$$\frac{[r(x)]^x}{r(\log x)} < \nu < \frac{1}{e}, \qquad m < x,$$

and is divergent if

$$\frac{[r(x)]^x}{r(\log x)} > \nu > \frac{1}{e}. \qquad m < x.$$

In case

$$\lim_{x=\infty} r(x) = 1,$$

this test is no more than the test of Schlömilch which may be stated as follows: The series converges if $x \log r(x) < \nu_1 < -1$, $m < x$; and diverges if $x \log r(x) > \nu_1 > -1$, $m < x$. From the point of view of the integral tests, Theorem IV may be used as a proof of Schlömilch's test.

3. An integral test involving $(r(x) - 1)$

The following is the most generally useful integral test of the second kind.

Theorem V. *Given the series*

$$u_0 + u_1 + u_2 + \cdots \qquad (u_n > 0,\ n > \mu).$$

Let $r_n = u_{n+1}/u_n$, and suppose that $r(x)$ is a positive, integrable function satisfying the preliminary conditions of one of the Theorems I, II, or III, and the further condition that

$$|r(x) - 1| < a < 1, \qquad \mu_1 < x.$$

Then a sufficient condition for the convergence of the series is the convergence of the integral

$$\int_{x_0}^{\infty} e^{\int_{x_0}^{x} (r(x) - 1)dx}\, dx.$$

This condition is also necessary for the convergence of the series if from a certain

point on

$$|r(x)-1|<\frac{k}{x},$$

where k is some constant.

Proof. We have the expansion

$$(1)\quad \log r(x)=(r(x)-1)-\tfrac{1}{2}(r(x)-1)^2+\tfrac{1}{3}(r(x)-1)^3-\cdots.$$

Then

$$\log r(x)\leqq r(x)-1,$$

so that

$$\int_{x_0}^{x} e^{\int_{x_0}^{x}\log r(x)dx}\,dx\leqq\int_{x_0}^{x} e^{\int_{x_0}^{x}(r(x)-1)dx}\,dx.$$

Therefore, since the preliminary conditions of one of the Theorems I, II, or III are satisfied, so that the given series and the integral

$$\int_{x_0}^{\infty} e^{\int_{x_0}^{x}\log r(x)dx}\,dx$$

converge or diverge together, the convergence of the integral

$$\int_{x_0}^{\infty} e^{\int_{x_0}^{x}(r(x)-1)dx}\,dx$$

is sufficient for the convergence of the given series.

The expansion (1) converges uniformly for $|r(x)-1|<a<1$, that is, for $\mu_1\leqq x$. We can therefore integrate it term by term over any interval $\mu_1\leqq x\leqq A$. Then

$$\int_{x_0}^{x}\log r(x)\,dx=\int_{x_0}^{x}(r(x)-1)\,dx-\frac{1}{2}\int_{x_0}^{x}(r(x)-1)^2\,dx+\cdots,$$

$$\mu_1<x_0<x,$$

and

$$\int_{x_0}^{x} e^{\int_{x_0}^{x}\log r(x)dx}\,dx$$

$$=\int_{x_0}^{x} e^{\int_{x_0}^{x}(r(x)-1)dx-\frac{1}{2}\int_{x_0}^{x}(r(x)-1)^2dx+\frac{1}{3}\int_{x_0}^{x}(r(x)-1)^3dx-\cdots}\,dx,\quad \mu_1<x_0<x.$$

We have

$$|r(x)-1|<\frac{k}{x},$$

so that

$$(r(x)-1)>-\frac{k}{x},\qquad -(r(x)-1)^2>-\frac{k^2}{x^2},$$

$$(r(x)-1)^3>-\frac{k^3}{x^3},\quad \text{etc.},\quad x>\mu_1.$$

Therefore

$$\int_{x_0}^{x} e^{\int_{x_0}^{x} \log r(x)dx}\,dx$$

$$> \int_{x_0}^{x} e^{\int_{x_0}^{x} (r(x)-1)dx - \frac{1}{2}\int_{x_0}^{x}\frac{k^2}{x^2}dx - \frac{1}{3}\int_{x_0}^{x}\frac{k^3}{x^3}dx - \frac{1}{4}\int_{x_0}^{x}\frac{k^4}{x^4}dx - \cdots}\,dx$$

$$> \int_{x_0}^{x} e^{\int_{x_0}^{x} (r(x)-1)dx - \frac{1}{2}\int_{x_0}^{\infty}\frac{k^2}{x^2}dx - \frac{1}{3}\int_{x_0}^{\infty}\frac{k^3}{x^3}dx - \frac{1}{4}\int_{x_0}^{\infty}\frac{k^4}{x^4}dx - \cdots}\,dx$$

$$= \int_{x_0}^{x} e^{\int_{x_0}^{x} (r(x)-1)dx - \left(\frac{k^2}{2x_0} + \frac{k^3}{3\cdot 2x_0^2} + \frac{k^4}{4\cdot 3x_0^3} + \cdots\right)}\,dx, \quad \mu_1 < x_0 < x.$$

If x_0 is taken greater than k, the series in parentheses converges to a value c_1. Consequently

$$\int_{x_0}^{x} e^{\int_{x_0}^{x} \log r(x)dx}\,dx > c_2 \int_{x_0}^{x} e^{\int_{x_0}^{x} (r(x)-1)dx}\,dx.$$

Therefore when x increases indefinitely, if the second of these two integrals diverges, the first one also diverges, and the given series diverges.

In most cases of interest $r(x)$ is a monotonically increasing function with unity as its limit. In such a case the condition $|r(x) - 1| < k/x$ becomes $(r(x) - 1) > -k/x$. This condition is not a very great restriction. For in case the integral

$$\int_{x_0}^{\infty} e^{\int_{x_0}^{x} (r(x)-1)dx}\,dx$$

diverges, if λ is any constant greater than unity the product of the integrand by λ^x is not finite for x infinite, and it is natural to assume that

$$x^{\lambda}\, e^{\int_{x_0}^{x} (r(x)-1)dx} > 1, \qquad m < x,$$

or

$$\int_{x_0}^{x} (r(x) - 1)\,dx > -\lambda \log x, \qquad m < x.$$

That is, if x_0 is properly chosen,

$$\int_{x_0}^{x} (r(x) - 1)\,dx > -\int_{x_0}^{x} \frac{\lambda}{x}\,dx, \qquad m < x_0 < x.$$

If this condition holds it is clear that for considerable intervals, at least,

$$r(x) - 1 > -\lambda/x.$$

So the condition that $r(x) - 1 > -k/x$, $m < x$, is a condition that we

may expect to hold rather commonly when the integral

$$\int_{x_0}^{\infty} e^{\int_{x_0}^{x}(r(x)-1)dx}\,dx$$

diverges.

As a corollary to the theorem we see that if $r(x)$ satisfies the preliminary conditions of the theorem, and if

$$|r(x)-1|<k/x, \qquad \mu \leqq x, \qquad k \leqq 1,$$

the series diverges. For

$$\int_{x_0}^{x} e^{\int_{x_0}^{x}(r(x)-1)dx}\,dx \geqq \int_{x_0}^{x} e^{-\int_{x_0}^{x}\frac{k}{x}dx}\,dx = \int_{x_0}^{x}\frac{c\,dx}{x^k}.$$

Similarly, if $r(x)-1<-k/x$, $k>1$, the series converges. In this way Raabe's test* may be proved.

The test of Theorem V is stronger than any test of the logarithmic scale, for by means of it we can test the series

$$\sum_{n=\mu}^{\infty}\frac{1}{n\cdot l_1 n\cdot l_2 n\cdot\cdots\cdot l_{\lambda-1}n\cdot(l_\lambda n)^p}.$$

For this series

$$r(x)=\frac{xl_1 xl_2 x\cdot\cdots\cdot l_{\lambda-1}x\,(l_\lambda x)^p}{(x+1)l_1(x+1)l_2(x+1)\cdot\cdots\cdot l_{\lambda-1}(x+1)\,(l_\lambda(x+1))^p}.$$

In all cases all of the preliminary conditions of the test hold for this series, except perhaps the condition that $|r(x)-1|<k/x$. Therefore if the integral of the test converges the series is known to converge. Suppose that the integral diverges. The series also will then diverge. For in case that it converged we should have

$$r(x)\geqq\left(\frac{x}{x+1}\right)^p=\left(1+\frac{1}{x}\right)^{-p};$$

and

$$r(x)-1>-2p/x, \qquad x>m,$$

which shows that all of the preliminary conditions of the theorem are then satisfied. Consequently, since the integral diverges, the theorem leads to a contradiction of the assumption that the series converges, or the integral and the series always converge or diverge together.

Just as we were able to establish d'Alembert's test and Raabe's first test, we can establish by means of Theorems I, II, III, and V most of the standard tests that use the test-ratio.

*Journal für Mathematik von Ettinghausen und Baumgartner, vol. 10; Journal de mathématiques, vol. 6 (1841), p. 85; Journal für Mathematik, vol. 11 (1832), p. 309.

Thus consider the series for which

$$r(x) = 1 - \frac{1}{x} - \frac{1}{xl_1x} - \cdots - \frac{1}{x \cdot l_1x \cdot \cdots \cdot l_{k-1}x} - \frac{p}{x \cdot l_1x \cdot \cdots \cdot l_kx}.$$

This series clearly satisfies the conditions for Theorem V.

$$\int_m^\infty e^{\int_m^x (r(x)-1)dx}\,dx$$

$$= \int_m^\infty e^{-\int_m^x \left[\frac{1}{x} + \frac{1}{xl_1x} + \cdots + \frac{1}{x \cdot l_1x \cdot \cdots \cdot l_{k-1}x} + \frac{p}{x \cdot l_1x \cdot \cdots \cdot l_kx}\right] dx}\,dx$$

$$= c\int_m^\infty \frac{dx}{xl_1xl_2x \cdots l_{k-1}x\,(l_kx)^p}.$$

Therefore the series converges if $p > 1$, and diverges if $p \leqq 1$.

We thus have the tests of de Morgan* and Bertrand†: *If $r(x)$ can be expressed in the form*

$$r(x) = 1 - \frac{1}{x} - \frac{1}{xl_1x} - \cdots - \frac{1}{xl_1x \cdots l_{k-1}x} - \frac{\omega(x)}{xl_1x \cdots l_kx}$$

and if

$$\lim_{x=\infty} \omega(x) = l,$$

the corresponding series converges when $l > 1$, and diverges when $l < 1$. In stating these tests it is more common to express the reciprocal of $r(x)$ as

$$\frac{1}{r(x)} = 1 + \frac{1}{x} + \frac{1}{xl_1x} + \cdots + \frac{1}{xl_1x \cdots l_{k-1}x} + \frac{\omega_1(x)}{xl_1x \cdots l_kx};$$

then *the series converges if*

$$\lim_{x=\infty} \omega_1(x) > 1,$$

and diverges if

$$\lim_{x=\infty} \omega_1(x) < 1.$$

The same tests are often stated in the following form. *If of the following limits*

$$\lim_{x=\infty} x\left(\frac{1}{r(x)} - 1\right) = k_0,$$

$$\lim_{x=\infty} l_1x\left(x\left(\frac{1}{r(x)} - 1\right) - 1\right) = k_1,$$

$$\lim_{x=\infty} l_2x\left(l_1(x)\left(x\left(\frac{1}{r(x)} - 1\right) - 1\right) - 1\right) = k_2, \quad \text{etc.},$$

* *Differential and Integral Calculus*, 1839.

† Journal de mathématiques, vol. 7 (1842), p. 37.

k_m is the first that is not equal to unity, the corresponding series converges if $k_m > 1$, and diverges if $k_m < 1$.

In his study of the hypergeometric series Gauss* gave the following rule. *If*

$$r_n = \frac{n^k + a_1 n^{k-1} + \cdots + a_k}{n^k + b_1 n^{k-1} + \cdots + b_k},$$

where k is a positive integer, the series converges if $(b_1 - a_1) > 1$, and diverges if $(b_1 - a_1) \leqq 1$. This test of course is merely the first test of the logarithmic scale. It is easily established directly by means of Theorem V, through one integration. This gives perhaps the easiest method of testing the hypergeometric series. The same results can be found by applying Theorem V directly to the hypergeometric series, though Gauss's test, so easily established by means of the theorem, is more convenient.

It is not difficult to show that in Theorem V the condition $|r(x) - 1| < k/x$ may be replaced by the condition that the series $\sum_{n=\mu}^{n=\infty}(1 - r_n)$ converges, and $r_n \leqq 1$. Then by the Maclaurin-Cauchy integral test, we see that *the series $\sum_{n=\mu}^{n=\infty} u_n$ diverges if the series $\sum_{n=\mu}^{n=\infty}(1 - r_n)$ converges and $r_n \leqq r_{n+1}$.* Of course this test is very weak. In the case of the series

$$1 + \frac{2^p - 1}{2^p} + \frac{(2^p - 1)(3^p - 1)}{(3!)^p} + \cdots,$$

for example, where

$$r_n = 1 - \frac{1}{n^p},$$

this test, like Raabe's first test, indicates divergence only for $p > 1$, though by Theorem V we see that the series also diverges for $p = 1$.

4. A series of products

We have noticed the normal form to which may be reduced any series capable of being tested by means of a test of the second kind. This normal form suggests a somewhat more general class of series. Suppose that we are given the two-dimensional array of numbers

$$\begin{array}{cccc} a_{00} & a_{01} & a_{02} & \cdots \\ a_{10} & a_{11} & a_{12} & \cdots \\ a_{20} & a_{21} & a_{22} & \cdots \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array},$$

and suppose that they form a series as follows:

$$a_{00} + a_{01} \cdot a_{10} + a_{02} \cdot a_{11} \cdot a_{20} + a_{03} \cdot a_{12} \cdot a_{21} \cdot a_{30} + \cdots.$$

* *Gesammelte Werke*, vol. 3, pp. 123–163.

The following theorem then holds

THEOREM VI. *Let $a(x, y)$ be a positive, integrable function for $x \geqq 0$, $y \geqq 0$, such that*

(*a*) $$a(m, n) = a_{m,\,n};$$

(*b*) $$0 < c < a(x, y) < k;$$

(*c*) *one of the two relations*

$$a(x', y') \leqq a(x', y''), \qquad 0 \leqq x', \qquad 0 \leqq y' < y''$$

or

$$a(x', y') \geqq a(x', y''), \qquad 0 \leqq x', \qquad 0 \leqq y' < y''$$

always holds; and

(*d*) *such that one of the two relations*

$$a(x', y') \leqq a(x'', y'), \qquad 0 \leqq x' < x'', \qquad 0 \leqq y'$$

or

$$a(x', y') \geqq a(x'', y'), \qquad 0 \leqq x' < x'', \qquad 0 \leqq y'$$

always holds.

Then a necessary and sufficient condition for the convergence of the series

$$\sum_{n=0}^{\infty} u_n \qquad \textit{where} \qquad u_n = \prod_{m=0}^{n} a_{m,\,n-m}$$

is the convergence of the integral

$$\int_0^{\infty} e^{\int_0^{\xi} \log a(x,\, \xi - x)\,dx}\, d\xi.$$

Proof. We can take $c < 1$ and $k > 1$. Take the case that $a(x', y)$ and $a(x, y')$ monotonically decrease as y and x respectively increase. Then

$$(1)\quad \log a(n+1, \xi - n) \leqq \int_n^{n+1} \log a(x, \xi - x)\,dx \leqq \log a(n, \xi - n - 1),$$

and

$$\int_n^{n+1} e^{\int_0^{\xi} \log a(x,\, \xi - x)\,dx}\, d\xi = \int_n^{n+1} e^{\overbrace{\int_0^1 + \int_1^2 + \cdots + \int_{n-1}^{n} + \int_n^{\xi}} \log a(x,\, \xi - x)\,dx}\, d\xi$$

$$\leqq \int_n^{n+1} e^{\log a(0,\, \xi - 1) + \log a(1,\, \xi - 2) + \cdots + \log a(n-1,\, \xi - n) + \log k}\, d\xi$$

$$\leqq k a_{0,\,n-1} \cdot a_{1,\,n-2} \cdot a_{2,\,n-3} \cdot \cdots \cdot a_{n-1,\,0}.$$

Therefore

$$(2)\qquad \int_n^{n+1} e^{\int_0^{\xi} \log a(x,\, \xi - x)\,dx}\, d\xi \leqq k u_{n-1}.$$

Likewise

$$\int_n^{n+1} e^{\int_0^\xi \log a(x,\xi-x)dx}\,d\xi$$

$$\geqq \int_n^{n+1} e^{\log a(1,\xi)+\log a(2,\xi-1)+\cdots+\log a(n,\xi-n+1)+\log C}\,d\xi$$

$$\geqq c a_{1,n+1}\cdot a_{2,n}\cdots a_{n,2} = \frac{c}{a_{0,n+2}\cdot a_{n+1,1}\cdot a_{n+2,0}}u_{n+2}.$$

Therefore

$$\int_n^{n+1} e^{\int_0^\xi \log a(x,\xi-x)dx}\,d\xi \geqq \frac{c}{k^3}u_{n+2}. \tag{3}$$

Similar inequalities hold if $a(x', y)$ and $a(x, y')$ increase monotonically, or if one decreases and the other increases.

The theorem then follows from a comparison of the given series with the series

$$\sum_{n=0}^{\infty}\int_n^{n+1} e^{\int_0^\xi \log a(x,\xi-x)dx}\,d\xi$$

by means of the inequalities (2) and (3).

The theorem can readily be extended to series for which the function $a(x, y)$ satisfies conditions similar to the conditions in Theorems II and III. A theorem analogous to Theorem V is also easily deduced.

As special cases this theorem in its extended form includes Theorems I, II, and III.

Example. The following series is conveniently tested by means of Theorem VI.

$$e^{-\beta/2-\alpha/2}+e^{-\beta/2-\beta/3-\alpha/3-\alpha/4}+e^{-\beta/2-\beta/3-\beta/4-\alpha/4-\alpha/6-\alpha/6}+\cdots.$$

This is formed from the array

$$\begin{array}{cccc} e^{-\beta/2-\alpha/2} & e^{-\beta/3-\alpha/3} & e^{-\beta/4-\alpha/4} & \cdots \\ e^{-\beta/2-\alpha/4} & e^{-\beta/3-\alpha/6} & e^{-\beta/4-\alpha/8} & \cdots \\ e^{-\beta/2-\alpha/6} & e^{-\beta/3-\alpha/9} & e^{-\beta/4-\alpha/12} & \cdots \\ \cdots & \cdots & \cdots & \cdots. \end{array}$$

Here

$$a_{m,n} = e^{-[\beta/(n+2)]-\alpha/[(m+1)(n+2)]},$$

and

$$a(x,y) = e^{-[\beta/(y+2)]-\alpha/[(x+1)(y+2)]},$$

where $0 \leqq x$, $0 \leqq y$. If we take the series as starting with the term $a_{1,1}$ instead of with $a_{0,0}$, we can write

$$a(x,y) = e^{-\beta/(y+1)-\alpha/(x(y+1))}, \qquad 1 \leqq x, \qquad 1 \leqq y.$$

Then

$$\int_1^\xi \log a(x, \xi - x)\,dx = \int_1^\xi \left[\frac{-\beta}{\xi + 1 - x} - \frac{\alpha}{(\xi+1)(\xi+1-x)} - \frac{\alpha}{(\xi+1)x} \right] dx$$

$$= -\left(\beta + \frac{2\alpha}{\xi+1}\right) \log \xi.$$

Consequently

$$\int_m^\xi e^{\int_1^\xi \log a(x, \xi - x)dx}\, d\xi = \int_m^\xi \frac{d\xi}{\xi^{\beta + [2\alpha/(\xi+1)]}}.$$

Therefore the series converges for $\beta > 1$, and diverges for $\beta \leqq 1$.

5. Multiple series

The tests given for simple series are easily generalized for multiple series. Thus for double series we have the following theorem:

Theorem VII. *Given the double series*

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} u_{m,n}, \qquad u_{m,n} > 0.$$

Let $r_m = u_{m+1,0}/u_{m,0}$, *and* $\rho_{m,n} = u_{m,n+1}/u_{m,n}$. *Suppose that* $r(x)$ *is a continuous function for* $0 \leqq x$, *having the properties that* $r(m) = r_m$, $0 < c < r(x) < k$, *and* $r(x)$ *monotonically increases as* x *increases; suppose also that* $\rho(x, y)$ *is a continuous function for* $0 \leqq x$, $0 \leqq y$, *having the properties that* $\rho(m, n) = \rho_{m,n}$, $0 < c < \rho(x, y) < k$, *and* $\rho(x, y)$ *increases monotonically when either* x *or* y *increases; then a necessary and sufficient condition for the convergence of the given double series is the convergence of the double integral*

$$\int_0^\infty \int_0^\infty e^{\int_0^x \log r(x)dx + \int_0^y \log \rho(x, y)dy}\, dx\, dy.$$

A proof like that of the first theorem can easily be given.

Example. Test the series

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} u_{m,n}$$

where $u_{00} = 1$, and

$$u_{m,n} = e^{-p - (p/2) - (p/3) - \cdots - [p/(m+n)]}, \qquad m + n > 0.$$

For this series we have

$$r_m = \frac{u_{m+1,0}}{u_{m,0}} = e^{-[p/(m+1)]}, \qquad \rho_{m,n} = \frac{u_{m,n+1}}{u_{m,n}} = e^{-[p/(m+n+1)]},$$

$$r(x) = e^{-[p/(x+1)]}, \qquad \rho(x, y) = e^{-[p/(x+y+1)]}.$$

Here we have

$$\int_0^x \log r(x)\,dx = \int_0^x \frac{-p}{x+1}\,dx = -p\log(x+1),$$

$$\int_0^y \log \rho(x,y)\,dy = \int_0^y \frac{-p}{x+y+1}\,dy = -p\log(x+y+1)+p\log(x+1),$$

and hence

$$\int_0^\infty \int_0^\infty e^{\int_0^x \log r(x)dx + \int_0^y \log \rho(x,y)dy}\,dx\,dy = \int_0^\infty \int_0^\infty \frac{dxdy}{(x+y+1)^p}$$

$$= \int_0^\infty \frac{dy}{(p-1)(y+1)^{p-1}}.$$

Therefore the series converges if $p > 2$, and diverges if $p \leqq 2$.

The theorems of this paper have been stated for constant terms. They can readily be extended to series of functions, not only to test the convergence of a given series, but also to determine whether the convergence is uniform; uniform convergence of an integral implies uniform convergence of the corresponding series.

CAMBRIDGE, MASS.

May, 1916

ON THE REDUCTION OF CERTAIN DIFFERENTIAL EQUATIONS OF THE SECOND ORDER*

BY

WILLIAM DUNCAN MACMILLAN

It is proposed to investigate in this paper the reduction of the differential equations of the second order

$$(1)\qquad x_1' = +p_1 x_1 + \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} a_{ij}\, x_1^i\, x_2^j = +p_1 x_1 + X_1,$$

$$x_2' = -p_2 x_2 + \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} b_{ij}\, x_1^i\, x_2^j = -p_2 x_2 + X_2 \qquad (i+j \geqq 2),$$

where p_1 and p_2 are two positive numbers, and the coefficients a_{ij}, b_{ij} are constants, to similar differential equations in which the right members are polynomials, by means of the linear-transcendental substitution

$$(2)\qquad x_1 = y_1 + \sum\sum c_{ij}\, y_1^i\, y_2^j,$$

$$x_2 = y_2 + \sum\sum d_{ij}\, y_1^i\, y_2^j \qquad (i+j \geqq 2).$$

Two essentially distinct cases present themselves: first, the ratio p_1/p_2 is irrational; second, the ratio p_1/p_2 is rational. It will be shown in the case $p_1/p_2 = \beta$ is irrational and the number β satisfies a rather mild condition that there exists a convergent transformation which reduces the differential equations to their linear terms. If the ratio is rational and the equations are canonical a convergent transformation exists for which the reduced differential equations have the form

$$y_1' = +y_1(1 + Q\, y_1\, y_2),$$

$$y_2' = -y_2(1 + Q\, y_1\, y_2),$$

which are easily integrated. If the ratio is rational and the equations are not canonical then it is always possible to reduce the differential equations to algebraic forms but nothing can be said as to the convergence of the transformations without further knowledge of the coefficients of equations (1),

* Presented to the Society, April 21, 1916.

except that divergent transformations are quite possible. A better statement, perhaps, would be that further knowledge of the coefficients is required before suitable algebraic forms can be determined. It seems certain that a knowledge of a few coefficients of terms of low degree is not sufficient.

Before entering upon the general discussion we shall show that it is no essential restriction upon the differential equations to suppose that x_1 is a factor of the right member of the first equation of (1) and that x_2 is a factor of the second equation; for if it were not so, and we take

$$\begin{aligned} x_1 &= y_1 + f_2(y_2), \\ x_2 &= y_2 + f_1(y_1) \end{aligned} \tag{3}$$

then $f_1(y_1)$ and $f_2(y_2)$ can be so chosen that the resulting equations in y_1 and y_2 will have this property. Making the substitution (3) in (1) and solving the resulting equations there is obtained

$$\begin{aligned} \left(1 - \frac{df_1}{dy_1}\frac{df_2}{dy_2}\right)y_1' &= +[p_1(y_1 + f_2) + X_1] + [p_2(y_2 + f_1) - X_2]\frac{df_2}{dy_2}, \\ \left(1 - \frac{df_1}{dy_1}\frac{df_2}{dy_2}\right)y_2' &= -[p_2(y_2 + f_1) - X_2] - [p_1(y_1 + f_2) + X_1]\frac{df_1}{dy_1}. \end{aligned} \tag{4}$$

The right members of these equations will carry y_1 and y_2 respectively as factors if the conditions

$$\begin{aligned} [p_1 y_1 + X_1(y_1, f_1)]\frac{df_1}{dy_1} + [p_2 f_1 - X_2(y_1, f_1)] &= 0, \\ [p_2 y_2 - X_2(f_2, y_2)]\frac{df_2}{dy_2} + [p_1 f_2 + X_1(f_2, y_2)] &= 0 \end{aligned} \tag{5}$$

are satisfied. These conditions are obtained by putting $y_1 = 0$ in the right member of the first equation and $y_2 = 0$ in the right member of the second equation. The two equations of (5) are of the same type, the first depending only on the variable y_1 and the second only on the variable y_2. The solution of either is dominated by the solution of

$$\left[y - \frac{M(y+f)^2}{1 - \alpha(y+f)}\right]\frac{df}{dy} + \beta f = \frac{M(y+f)^2}{1 - \alpha(y+f)} \tag{5'}$$

where M and α are positive constants suitably chosen, and $\beta = p_1/p_2$ or its reciprocal. If we take $f = y\phi$ it is found that the solution of (5′) is dominated by the solution of

$$y\frac{\partial\phi}{\partial y} + (1+\beta)\phi = \frac{My(1+\beta\phi)(1+\phi)^2}{1 - \alpha y(1+\phi) - My(1+\phi)^2}, \tag{5''}$$

and the solution of (5″) is dominated by the solution of

$$(1+\beta)\phi = \frac{My(1+\beta\phi)(1+\phi)^2}{1-\alpha y(1+\phi) - My(1+\phi)^2}$$

which is convergent, whatever may be the value of β, since it is always positive.

If therefore f_1 and f_2 satisfy conditions (5) equations (4) are of the same form as (1) except that y_1 is a factor of the right member of the first equation and y_2 is a factor of the right member of the second. It is therefore no essential restriction to assume that the differential equations are of this form.

I. The ratio p_1/p_2 is irrational

Let us suppose that p_1/p_2 is irrational and equal to γ. Then the equations can be written

$$\begin{aligned} x_1' &= x_1(+1+X_1), \\ x_2' &= x_2(-\gamma+X_2), \end{aligned} \tag{6}$$

where X_1 and X_2 are power series in x_1 and x_2 of order one. We make now the substitution

$$x_1 = y_1(1+f_1) = y_1 e^{g_1}, \qquad x_2 = y_2(1+f_2) = y_2 e^{g_2}, \tag{7}$$

where f_1, f_2, g_1, g_2 are power series in y_1 and y_2 vanishing for $y_1 = y_2 = 0$. If we assume that

$$y_1' = y_1, \qquad y_2' = -\gamma y_2, \tag{8}$$

we obtain

$$\begin{aligned} y_1\frac{\partial g_1}{\partial y_1} - \gamma y_2\frac{\partial g_1}{\partial y_2} &= X_1(y_1 e^{g_1}, y_2 e^{g_2}), \\ y_1\frac{\partial g_2}{\partial y_2} - \gamma y_2\frac{\partial g_2}{\partial y_2} &= X_2(y_1 e^{g_1}, y_2 e^{g_2}). \end{aligned} \tag{9}$$

If we set

$$\begin{aligned} g_1 &= \sum\sum g_{ij}^{(1)} y_1^i y_2^j, \\ g_2 &= \sum\sum g_{ij}^{(2)} y_1^i y_2^j, \end{aligned} \tag{10}$$

we find for the determination of the coefficients $g_{ij}^{(1)}$ and $g_{ij}^{(2)}$ the relations

$$\begin{aligned} (i-j\gamma)g_{ij}^{(1)} &= G_{ij}^{(1)}, \\ (i-j\gamma)g_{ij}^{(2)} &= G_{ij}^{(2)}, \end{aligned}$$

where at each step $G_{ij}^{(1)}$ and $G_{ij}^{(2)}$ are known quantities. Since γ is irrational none of the coefficients $(i-j\gamma)$ vanishes and there is no difficulty in determining the coefficients $g_{ij}^{(1)}$ and $g_{ij}^{(2)}$ which define the transformation.

It follows from (9) and the fact that g_1 and g_2 are power series in y_1 and y_2

of order one that $g_1 = g_2$ if $X_1 = X_2$. For the question of convergence therefore it will be sufficient to consider the single equation

$$y_1 \frac{\partial g}{\partial y_1} - \gamma y_2 \frac{\partial g}{\partial y_2} = \left[\frac{1}{1-y_1} \cdot \frac{1}{1-y_2} - 1 \right] \frac{M}{1-g}, \tag{11}$$

which dominates (9). Since γ occurs with the negative sign in the left member of this equation the solution does not have all of its signs positive and one could not be sure that the solution of (11) dominates the solution of (9). Let us consider first the solution of

$$y_1 \frac{\partial g}{\partial y_1} + \gamma y_2 \frac{\partial g}{\partial y_2} = \left[\frac{1}{1-y_1} \cdot \frac{1}{1-y_2} - 1 \right] \frac{M}{1-g}. \tag{12}$$

The solution of this equation has all of its terms positive and no cancellation occurs, and the coefficients so far as they depend upon the right member of (12) have maximum values. The solution of (12) is easily obtained, for if we take

$$\omega = g - \tfrac{1}{2}g^2$$

we get

$$y_1 \frac{\partial \omega}{\partial y_1} + \gamma y_2 \frac{\partial \omega}{\partial y_2} = M \sum \sum y_1^i y_2^j$$

and consequently

$$\omega = M \sum \sum \frac{y_1^i y_2^j}{i + j\gamma}.$$

Since $g = 1 - \sqrt{1 - 2\omega}$ is expansible in powers of ω with positive coefficients it is easy to see how the function g is built up and all of the coefficients so far as they depend upon the right member of (12) have maximum values.

If now we change γ into $-\gamma$ the solution of (12) becomes the solution of (11). We have

$$g = 1 - \sqrt{1 - 2\omega},$$

$$\omega = M \sum \sum \frac{y_1^i y_2^j}{i - j\gamma}.$$

Since the series

$$\sum \sum \frac{y_1^i y_2^j}{|i - j\gamma|}$$

is convergent for a certain class of irrational numbers*, γ, it follows that if we take

$$\omega^* = M \sum \sum \frac{y_1^i y_2^j}{|i - j\gamma|}, \qquad g^* = 1 - \sqrt{1 - 2\omega^*}$$

*See Bulletin of the American Mathematical Society, vol. 22 (1915), p. 26–32.

then g^* will be a convergent series which will dominate the solutions (9) for the coefficients of g^* will be a maximum in so far as they depend upon the right members through M and a maximum in so far as they depend upon γ being negative.

It follows therefore, since $y_1 = \mu_1 e^t$, $y_2 = \mu_2 e^{-\gamma t}$, that x_1 and x_2 are expansible in powers of $\mu_1 e^t$ and $\mu_2 e^{-\gamma t}$, and that these expansions are convergent for any preassigned value of t provided μ_1 and μ_2 are sufficiently small.

II. The ratio p_1/p_2 is rational

If the ratio p_1/p_2 is rational it is always possible by a linear change of the independent variable, t, to make p_1 and p_2 integers, and we will suppose this to have been done. Then by the substitution

$$x_1 = \xi_1^{p_1}, \qquad x_2 = \xi_2^{p_2},$$

we will have, since

$$\frac{1}{p_1}\frac{x_1'}{x_1} = \frac{\xi_1'}{\xi_1}, \qquad \frac{1}{p_2}\frac{x_2'}{x_2} = \frac{\xi_2'}{\xi_2},$$

two equations in ξ_1 and ξ_2 of the same type as (1), in which p_1 and p_2 are each unity. We may therefore suppose $p_1 = p_2 = 1$, and the equations are

$$\begin{aligned} x_1' &= x_1(+1 + X_1^{(1)}), \\ x_2' &= x_2(-1 + X_2^{(2)}). \end{aligned} \tag{13}$$

For our present purposes, however, it will be preferable to use the less specialized form

$$\begin{aligned} x_1' &= +x_1 + X_1, \\ x_2' &= -x_2 + X_2, \end{aligned} \tag{13}$$

where

$$X_1 = \sum\sum a_{ij} x_1^i x_2^j, \qquad X_2 = \sum\sum b_{ij} x_1^i x_2^j \qquad (i+j \geqq 2),$$

in which it is not supposed that x_1 is a factor of X_1 nor x_2 a factor of X_2.

If, in (13), we make the linear-transcendental substitution

$$\begin{aligned} x_1 &= y_1 + \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} c_{ij} y_1^i y_2^j = y_1 + f_1(y_1, y_2), \\ x_2 &= y_2 + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} d_{ij} y_1^i y_2^j = y_2 + f_2(y_1, y_2) \end{aligned} \qquad (i+j \geqq 2), \tag{14}$$

it is found at once that the differential equations cannot be reduced to their linear terms, for it is not possible, in general, to determine the constants c_{ij} and d_{ij} so as to do this. It is, however, always possible to reduce them to the form

$$
\begin{aligned}
y_1' &= +y_1 + Q_1^{(1)} y_1^2 y_2 + Q_1^{(2)} y_1^3 y_2^2, \\
y_2' &= -y_2 + Q_2^{(1)} y_1 y_2^2 + Q_2^{(2)} y_1^2 y_2^3,
\end{aligned} \tag{15}
$$

in which $Q_i^{(j)}$ are constants provided $Q_1^{(1)} + Q_2^{(1)} \neq 0$. On substituting (14) and (15) in (13) there results

$$
\begin{aligned}
\sum_{i,j} (i-j-1)\, c_{ij}\, y_1^i\, y_2^j = \sum a_{ij} (y_1+f_1)^i (y_2+f_2)^j &- Q_1^{(1)} y_1^2 y_2 \\
&- Q_1^{(2)} y_1^3 y_2^2 - [Q_1^{(1)} y_1^2 y_2 + Q_1^{(2)} y_1^3 y_2^2] \frac{\partial f_1}{\partial y_1} \\
&- [Q_2^{(1)} y_1 y_2^2 + Q_2^{(2)} y_1^2 y_2^3] \frac{\partial f_1}{\partial y_2}, \\
\sum (i-j+1)\, d_{ij}\, y_1^i\, y_2^j = \sum b_{ij} (y_1+f_1)^i (y_2+f_2)^j &- Q_2^{(1)} y_1 y_2^2 \\
&- Q_2^{(2)} y_1^2 y_2^3 - [Q_1^{(1)} y_1^2 y_2 + Q_1^{(2)} y_1^3 y_2^2] \frac{\partial f_2}{\partial y_1} \\
&- [Q_2^{(1)} y_1 y_2^2 + Q_2^{(2)} y_1^2 y_2^3] \frac{\partial f_2}{\partial y_2}.
\end{aligned} \tag{16}
$$

From these equations it is seen that the c_{ij} are readily determined except for such values of i and j that $i - j = +1$ and that the d_{ij} are readily determined except when $i - j = -1$. Thus, up to and including terms of the second degree,

$$
\begin{aligned}
x_1 &= y_1 + [a_{20} y_1^2 - a_{11} y_1 y_2 - \tfrac{1}{3} a_{02} y_2^2], \\
x_2 &= y_2 + [\tfrac{1}{3} b_{20} y_1^2 + b_{11} y_1 y_2 - b_{02} y_2^2].
\end{aligned}
$$

Using these values it is found that the terms of third degree in the right members of (16) become

$$
\begin{aligned}
R_1^{(3)} = {} & (2a_{20}^2 + \tfrac{1}{3} a_{11} b_{20} + a_{30}) y_1^3 \\
& + (-a_{20} a_{11} + a_{11} b_{11} + \tfrac{2}{3} a_{02} b_{20} + a_{21} - Q_1^{(1)}) y_1^2 y_2 \\
& + (-\tfrac{2}{3} a_{20} a_{02} - a_{11}^2 - a_{11} b_{02} + 2a_{02} b_{11} + a_{12}) y_1 y_2^2 \\
& + (-\tfrac{1}{3} a_{11} a_{02} - 2a_{02} b_{02} + a_{03}) y_1^3, \\
R_2^{(3)} = {} & (2a_{20} b_{20} + \tfrac{1}{3} b_{20} b_{11} + b_{30}) y_1^3 \\
& + (-2b_{20} a_{11} + b_{11} a_{20} + \tfrac{2}{3} b_{20} b_{02} + b_{21}) y_1^2 y_2 \\
& + (-\tfrac{2}{3} b_{20} a_{02} - a_{11} b_{11} + b_{11} b_{02} + b_{12} - Q_2^{(1)}) y_1 y_2^2 \\
& + (-\tfrac{1}{3} b_{11} a_{02} - 2b_{02}^2 + b_{03}) y_2^3.
\end{aligned}
$$

Since in the left members of (16) the coefficients of c_{21} and d_{12} are zero these quantities cannot be determined and they therefore remain arbitrary. From the right members therefore we must have

$$
\begin{aligned}
Q_1^{(1)} &= -a_{20}\,a_{11} + a_{11}\,b_{11} + \tfrac{2}{3}a_{02}\,b_{20} + a_{21}, \\
Q_2^{(1)} &= -\tfrac{2}{3}b_{20}\,a_{02} - a_{11}\,b_{11} + b_{11}\,b_{02} + b_{12}, \\
Q_1^{(1)} + Q_2^{(1)} &= -a_{20}\,a_{11} + b_{11}\,b_{02} + a_{21} + b_{12},
\end{aligned} \tag{17}
$$

and then we obtain

$$
\begin{aligned}
c_{30} &= a_{20}^2 + \tfrac{1}{6}a_{11}\,b_{20} + \tfrac{1}{2}a_{30}, \\
c_{21} &= \alpha_1 = \text{arbitrary}, \\
c_{12} &= \tfrac{1}{3}a_{20}\,a_{02} + \tfrac{1}{2}a_{11}^2 + \tfrac{1}{2}a_{11}\,b_{02} - a_{02}\,b_{11} - \tfrac{1}{2}a_{12}, \\
c_{03} &= \tfrac{1}{12}a_{11}\,a_{02} + \tfrac{1}{2}a_{02}\,b_{02} - \tfrac{1}{4}a_{03}, \\
d_{30} &= \tfrac{1}{2}a_{20}\,b_{20} + \tfrac{1}{12}b_{20}\,b_{11} + \tfrac{1}{4}b_{30}, \\
d_{21} &= -b_{20}\,a_{11} + \tfrac{1}{2}b_{11}\,a_{20} + \tfrac{1}{2}b_{11}^2 + \tfrac{1}{3}b_{20}\,b_{02} + \tfrac{1}{2}b_{21}, \\
d_{12} &= \beta_1 = \text{arbitrary}, \\
d_{03} &= \tfrac{1}{6}b_{11}\,a_{02} + b_{02}^2 - \tfrac{1}{2}b_{03}.
\end{aligned}
$$

One can continue the computation as far as is desired except for the terms $y_1^{k+1}\,y_2^k$ in the first equation and $y_1^k\,y_2^{k+1}$ in the second. The coefficients $c_{k+1,\,k}$ $d_{k,\,k+1}$ (which for brevity we shall denote by α_k and β_k) cannot be determined as are the remaining coefficients. These coefficients which are left arbitrary in the terms of degree $2k+1$ enter also in terms of higher degree and are determined by the condition that in the right members of (16) the coefficients of the terms $y_1^{k+2}\,y_2^{k+1}$ in the first equation and $y_1^{k+1}\,y_2^{k+2}$ in the second equation must vanish. Thus it is found from the coefficients of $y_1^3\,y_2^2$ and $y_1^2\,y_2^3$ that

$$
\begin{aligned}
Q_1^{(2)} &= -Q_2^{(1)}\,\alpha_1 + Q_1^{(1)}\,\beta_1 + \text{known terms}, \\
Q_2^{(2)} &= +Q_2^{(1)}\,\alpha_1 - Q_1^{(1)}\,\beta_1 + \text{known terms}.
\end{aligned} \tag{18}
$$

Since the determinant of the coefficients of α_1 and β_1 in these two expressions is zero, α_1 and β_1 cannot in general be chosen so as to make $Q_1^{(2)} = 0$, $Q_2^{(2)} = 0$. It is necessary therefore to retain these constants in the differential equations. But $k = 1$ happens to be exceptional. For a general value of k these two conditions are

$$
\begin{aligned}
-[(k-1)\,Q_1^{(1)} + kQ_2^{(1)}]\,\alpha_k + Q_1^{(1)}\,\beta_k &= \text{known terms}, \\
Q_2^{(1)}\,\alpha_k - [kQ_1^{(1)} + (k-1)\,Q_2^{(1)}]\,\beta_k &= \text{known terms}.
\end{aligned} \tag{18'}
$$

The value of the determinant $D^{(k)}$ of the coefficients of α_k and β_k is

$$
D^{(k)} = k(k-1)\,(Q_1^{(1)} + Q_2^{(1)})^2.
$$

If therefore $Q_1^{(1)} + Q_2^{(1)} \neq 0$ the determinant vanishes only if $k = 0$ or $k = 1$, and there is no further difficulty in the determination of the coefficients. The determination is unique aside from the fact that α_1 and β_1 remain arbitrary. Since α_1 and β_1 enter $Q_1^{(2)}$ and $Q_2^{(2)}$ linearly, as is seen from (18), it will be convenient to impose the relations

$$Q_1^{(2)} = mQ_1^{(1)}, \qquad Q_2^{(2)} = mQ_2^{(1)},$$

which is always possible if $Q_1^{(1)} + Q_2^{(1)} \neq 0$, for on adding the two equations of (18) it is found that m equals the sum of the known terms divided by $Q_1^{(1)} + Q_2^{(1)}$. The quantity m is independent of α_1 and β_1 and depends only upon the coefficients of the original differential equations.

With these determinations therefore the differential equations take the form

$$\begin{aligned} y_1' &= y_1[+1 + Q_1^{(1)}(1 + my_1y_2)y_1y_2], \\ y_2' &= y_2[-1 + Q_2^{(1)}(1 + my_1y_2)y_1y_2]. \end{aligned} \tag{19}$$

which can be integrated. On taking

$$y_1 = z_1 e^t, \qquad y_2 = z_2 e^{-t},$$

the differential equations become

$$\begin{aligned} z_1' &= Q_1^{(1)} z_1[z_1z_2 + mz_1^2z_2^2], \\ z_2' &= Q_2^{(1)} z_2[z_1z_2 + mz_1^2z_2^2]. \end{aligned} \tag{20}$$

From these equations we get

$$(z_1z_2)' = (Q_1^{(1)} + Q_2^{(1)})z_1z_2[z_1z_2 + mz_1^2z_2^2]. \tag{21}$$

whence

$$\frac{1}{Q_1^{(1)}}\frac{z_1'}{z_1} = \frac{1}{Q_2^{(1)}}\frac{z_2'}{z_2} = \frac{1}{(Q_1^{(1)} + Q_2^{(1)})}\frac{(z_1z_2)'}{(z_1z_2)},$$

so that

$$z_1 = c(z_1z_2)^{Q_1^{(1)}/(Q_1^{(1)}+Q_2^{(1)})}, \qquad z_2 = \frac{1}{c}(z_1z_2)^{Q_2^{(1)}/(Q_1^{(1)}+Q_2^{(1)})},$$

$$z_1^{Q_2^{(1)}} = c^2 z_2^{Q_1^{(1)}}.$$

Using the function $L'(\tau)$ which was introduced in the discussion of the differential equation of the first order* it is seen from (21) that, if $m \neq 0$,

$$z_1z_2 = \frac{L'(\tau) - 1}{m},$$

where

$$\tau = 1 - \frac{Q_1^{(1)} + Q_2^{(1)}}{m}(t - t_0).$$

*See Annals of Mathematics, vol. 19 (1917), p. 21–29.

But if $m = 0$ we have

$$z_1 z_2 = \frac{1}{(Q_1^{(1)} + Q_2^{(1)})(t_0 - t)}.$$

Consequently, if $m \neq 0$, the expressions for y_1 and y_2 are

$$y_1 = ce^t \left(\frac{L'(\tau) - 1}{m} \right)^{Q_1^{(1)}/(Q_1^{(1)}+Q_2^{(1)})},$$

$$y_2 = \frac{1}{c} e^{-t} \left(\frac{L'(\tau) - 1}{m} \right)^{Q_2^{(1)}/(Q_1^{(1)}+Q_2^{(1)})}. \tag{22}$$

But if $m = 0$ then

$$y_1 = \frac{ce^t}{[(Q_1^{(1)} + Q_2^{(1)})(t_0 - t)]^{Q_1^{(1)}/(Q_1^{(1)}+Q_2^{(1)})}},$$

$$y_2 = \frac{\frac{1}{c} e^{-t}}{[(Q_1^{(1)} + Q_2^{(1)})(t_0 - t)]^{Q_2^{(1)}/(Q_1^{(1)}+Q_2^{(1)})}}. \tag{23}$$

It is to be observed also that equations (19) admit the integral

$$H = y_1^{-(Q_2^{(1)}+m)} y_2^{(Q_1^{(1)}-m)} (1 + m y_1 y_2)^m e^{-1/y_1 y_2}, \tag{24}$$

or, in logarithmic form,

$$H_1 = (m + Q_2^{(1)}) \log y_1 + (m - Q_1^{(1)}) \log y_2 - m \log (1 + m y_1 y_2) + \frac{1}{y_1 y_2}. \tag{25}$$

The following example is illustrative of this transformation

$$x_1' = x_1 + x_1^2 + x_1 x_2, \qquad x_2' = -x_2. \tag{26}$$

The transformation defined by the relations

$$x_1 = \frac{y_1 e^{-y_2}}{1 + y_1 y_2 \int \frac{e^{-y_2} + y_2}{y_2^2} dy_2}, \qquad x_2 = y_2, \tag{27}$$

transforms the differential equations into

$$y_1' = y_1(1 - y_1 y_2), \qquad y_2' = -y_2,$$

which have the form of (19) in which $Q_1^{(1)} = -1$, $m = Q_2^{(1)} = 0$, and it will be observed that the transformation considered as a power series in y_1 and y_2 is convergent.

It is not necessary however that the reduced equations shall have the form of (19). Indeed, it is evident from the nature of the process that any desired

terms, of degree not less than two, could be added to the right members of (19) and the process of determining the transformation would not fail. Thus it is possible to have an infinite variety of transformations and reduced differential equations, which shows clearly that all of these transformations cannot be convergent. Consider for example, the equations

$$(28)\qquad \begin{aligned} x_1' &= x_1[+1+(1+\alpha i_1)x_1x_2+\mu x_1^{i_2+1}x^{i_1+1}],\\ x_2' &= x_2[-1-(1+\alpha i_2)x_1x_2-\mu x_1^{i_2+1}x^{i_1+1}], \end{aligned}$$

in which α and μ are any constants and i_1, i_2 are positive integers or zero, but not equal. Here $Q_1^{(1)}+Q_2^{(1)}=\alpha(i_1-i_2)$ which is not zero if α is not zero. Consequently one can determine a transformation which will reduce (28) to (19). Equations (28) however admit the integral

$$(29)\qquad I = x_1^{-i_2}x_2^{-i_1}se^{s^a/a}+\mu\int e^{s^a/a}\,ds,$$

where $s^a=1/x_1x_2$. If now $\mu\neq 0$ the singularities of I in the neighborhood of the origin are not the same as the singularities of (24). Since $I(y_1y_2)$ is not a function of $H(y_1y_2)$, their jacobian being distinct from zero, it is clear that I cannot be transformed into H by a convergent linear-transcendental substitution, which cannot alter the character of the singularities at the origin.* We conclude therefore that the transformation in this case is divergent if μ is distinct from zero. Furthermore, since i_1 and i_2 may be as large as we please we must conclude that the singularities of the integral in the neighborhood of the origin do not depend merely upon the coefficients of the terms of low degree in the differential equations. It would seem to be a reasonable conjecture that, if the differential equations admit an integral which is known, then there exists a convergent linear-transcendental transformation which reduces the differential equations to an algebraic type such that the reduced equations admit an integral having the same singularities at the origin as the original integral, but this is not always true, as is shown by the following example. The equations

$$(A)\qquad \begin{aligned} y_1' &= \frac{y_1}{1+y_1}(+1+y_1-y_1y_2),\\ y_2' &= \frac{y_2}{1+y_1}(-1-y_1-y_1y_2) \end{aligned}$$

admit the integral

$$H_1 = y_2se^{s^2/2}+\int e^{s^2/2}\,ds, \qquad \text{where} \qquad s^2=\frac{1}{y_1y_2}.$$

* Dulac has considered an example in which he has shown from the law of the coefficients that the transformation in that case is divergent. See Bulletin des sciences mathématiques, vol. 37 (1913).

Likewise the equations

$$(B)\qquad \begin{aligned} x_1' &= x_1(+1+x_1-x_1x_2),\\ x_2' &= x_2(-1-x_1-x_1x_2) \end{aligned}$$

admit the integral

$$H_2 = x_2\sigma e^{\sigma^2/2} + \int e^{\sigma^2/2}\,d\sigma, \qquad \text{where} \qquad \sigma^2 = \frac{1}{x_1x_2}.$$

But (A) cannot be transformed into (B) by a linear-transcendental substitution, for we have

$$\frac{y_1'}{y_1} - \frac{y_2'}{y_2} = 2, \qquad \frac{(x_1x_2)'}{(x_1x_2)^2} = -2 \qquad \text{so that} \qquad \frac{y_1'}{y_1} - \frac{y_2'}{y_2} = -\frac{(x_1x_2)'}{(x_1x_2)^2}.$$

That is $y_1/y_2 = Ce^{1/x_1x_2}$. Obviously y_1/y_2 cannot be transformed into Ce^{1/x_1x_2} by a linear transcendental substitution.

EXISTENCE OF AN INTEGRAL OF THE FORM $H = \dfrac{P_1(x_1, x_2)}{P_2(x_1, x_2)}$

It will be shown in the present section, that if there exists an integral of the form

$$H = \frac{P_1(x_1, x_2)}{P_2(x_1, x_2)},$$

where P_1 and P_2 are ordinary power series (and this hypothesis includes the case where the differential equations are canonical) then the transformation is convergent and the reduced equations are

$$(30)\qquad \begin{aligned} y_1' &= y_1(+1+Qy_1y_2),\\ y_2' &= y_2(-1-Qy_1y_2). \end{aligned}$$

We have already observed that it is always possible to reduce the differential equations to the form

$$(31)\qquad \begin{aligned} y_1' &= y_1(+1+Q_1^{(1)}y_1y_2 + Q_1^{(1)}y_1^2y_2^2 + Q_1^{(3)}y_1^3y_2^3 + \cdots),\\ y_2' &= y_2(-1+Q_2^{(1)}y_1y_2 + Q_2^{(2)}y_1^2y_2^2 + Q_2^{(3)}y_1^3y_2^3 + \cdots), \end{aligned}$$

whether $Q_1^{(1)} + Q_2^{(1)} = 0$ or not zero, the $Q_i^{(j)}$ being suitably chosen constants.

On differentiating $H = P_1/P_2$ we have

$$(32)\qquad P_2\left[\frac{\partial P_1}{\partial x_1}x_1' + \frac{\partial P_1}{\partial x_2}x_2'\right] - P_1\left[\frac{\partial P_2}{\partial x_1}x_1' + \frac{\partial P_2}{\partial x_2}x_2'\right] \equiv 0.$$

Let us suppose now that $P_1 = p_1^{(n)} + p_1^{(n+1)} + \cdots$, $P_2 = p_2^{(m)} + p_2^{(m+1)} + \cdots$, where $p_j^{(k)}$ is a polynomial homogeneous in x_1 and x_2 of degree k. Consider the terms of lowest degree, $m+n$, in (32). We have

$$(33) \qquad p_2^{(m)}\left(x_1\frac{\partial p_1^{(n)}}{\partial x_1}-x_2\frac{\partial p_1^{(n)}}{\partial x_2}\right)-p_1^{(n)}\left(x_1\frac{\partial p_2^{(m)}}{\partial x_1}-x_2\frac{\partial p_2^{(m)}}{\partial x_2}\right)\equiv 0.$$

Take

$$p_1^{(n)}=\sum_{j=0}^{n}a_j\,x_1^j\,x_2^{n-j}, \qquad p_2^{(m)}=\sum_{k=0}^{m}b_k\,x_1^k\,x_2^{m-k}.$$

Then it follows from (33) that

$$(34) \qquad \sum_{j=0}^{n}\sum_{k=0}^{m}[\,2(j-k)-(n-m)\,]\,a_j\,b_k\,x_1^{j+k}\,x^{(m+n)-(j+k)}\equiv 0.$$

Since not every $a_j=0$ and not every $b_k=0$ we can suppose that $a_{j_1}\neq 0$, $b_{k_1}\neq 0$. Hence we must have

$$2(j_1-k_1)-(n-m)=0,$$

and consequently $n-m$ is an even integer, say $m=n-2r$. For any other value of j, say $j_s\neq j_1$ we have

$$2(j_s-k_1)-(n-m)\neq 0$$

and therefore $a_{j_s}=0$. Likewise if $k_s\neq k_1$ we have also

$$2(j_1-k_s)-(n-m)\neq 0$$

and therefore $b_{k_s}=0$. Hence $p_1^{(n)}$ and $p_2^{(m)}$ have only one term each, and without loss of generality the coefficients of these terms can each be taken equal to unity. Furthermore we can suppose $m\neq n$. For if $m=n$ we could take the integral

$$H_1=H-1=\frac{P_1}{P_2}-1=\frac{P_1-P_2}{P_2},$$

for which $m\neq n$, and consequently $r\neq 0$.

We have then

$$p_1^{(n)}=x_1^{j_1}\,x_2^{n-j_1}, \qquad p_2^{(m)}=x^{j_1}\,x^{(n-j_1)}\,(x_1\,x_2)^{-r}.$$

Let us now make the transformation

$$x_1=y_1+f_1(y_1,y_2), \qquad x_2=y_2+f_2(y_1,y_2),$$

such that the transformed differential equations are (31). Then $P_1(x_1,x_2)$ becomes

$$S_1(y_1,y_2)=y_1^{j_1}\,y_2^{n-j_1}+s_1^{(n+1)}+s_1^{(n+2)}+\cdots,$$

and $P_2(x_1,x_2)$ becomes

$$S_2(y_1,y_2)=y_1^{j_1}\,y_2^{n-j_1}\,(y_1\,y_2)^{-r}+s_2^{(m+1)}+s_2^{(m+2)}+\cdots.$$

On differentiating

$$H = \frac{S_1(y_1 y_2)}{S_2(y_1 y_2)}$$

we have

$$S_2\left[\frac{\partial S_1}{\partial y_1}y_1' + \frac{\partial S_1}{\partial y_2}y_2'\right] - S_1\left[\frac{\partial S_2}{\partial y_1}y_1' + \frac{\partial S_2}{\partial y_2}y_2'\right] \equiv 0. \tag{35}$$

Consider the terms of degree $n + m + 1$ in the expression of (35),

$$s_2^{(m+1)}\left(y_1\frac{\partial s_1^{(n)}}{\partial y_1} - y_2\frac{\partial s_1^{(n)}}{\partial y_2}\right) - s_1^{(n+1)}\left(y_1\frac{\partial s_2^{(m)}}{\partial y_1} - y_2\frac{\partial s_2^{(m)}}{\partial y_2}\right) \equiv 0.$$

On taking

$$s_1^{(n+1)} = \sum_{k=0}^{n+1} c_k\, y_1^k\, y_2^{n+1-k}, \qquad s_2^{(m+1)} = \sum_{j=0}^{m+1} d_j\, y_1^j\, y_2^{m+1-j}$$

we find, after removing a common factor,

$$\sum_{k=0}^{n+1}[2(k-j_1)-1]\,c_k\,y_1^{k-r}\,y_2^{n+1-r-k} - \sum_{j=1}^{n+1-2r}[2(j-j_1)+2r-1]\,d_j\,y_1^j\,y^{n+1-j-2r}.$$

If r is positive the first $\sum$ contains more terms than the second, and if r is negative the second contains more than the first. Let us suppose $r > 0$ (the result is the same for $r < 0$). Then $c_k = 0$ for $k = 0, \cdots, r-1$, $n+2-r, \cdots, n+1$. If now we take $k - r = j$ for the remaining terms we get

$$\sum_{j=0}^{n+1-2r}[2(j-j_1+r)-1][c_j - d_j]\,y_1^j\,y_2^{n+1-2r-j} \equiv 0,$$

and therefore $c_j = d_j$, since $2(j - j_1 + r) - 1$ is an odd integer. Hence

$$s_1^{(n+1)} = (y_1 y_2)^r\, s_2^{(m+1)}. \tag{36}$$

From the terms of degree $m + n + 2$ we obtain

$$\begin{aligned}
&s_2^{(m)}\left(y_1\frac{\partial s_1^{(n+2)}}{\partial y_1} - y_2\frac{\partial s_1^{(n+2)}}{\partial y_2}\right) - s_1^{(n)}\left(y_1\frac{\partial s_2^{(m+2)}}{\partial y_2} - y_2\frac{\partial s_2^{(m+2)}}{\partial y_2}\right)\\
&+ s_2^{(m+1)}\left(y_1\frac{\partial s^{(n+1)}}{\partial y_1} - y_2\frac{\partial s^{(n+1)}}{\partial y_2}\right) - s_1^{(n+1)}\left(y_1\frac{\partial s_2^{(m+1)}}{\partial y_1} - y_2\frac{\partial s_2^{(m+1)}}{\partial y_2}\right)\\
&+ s_2^{(m+2)}\left(y_1\frac{\partial s_1^{(n)}}{\partial y_1} - y_2\frac{\partial s_1^{(n)}}{\partial y_2}\right) - s_1^{(n+2)}\left(y_1\frac{\partial s_2^{(m)}}{\partial y_1} - y_2\frac{\partial s_2^{(m)}}{\partial y_2}\right)\\
&+ r(Q_1^{(1)} + Q_2^{(1)})\,y_1^{2j_1-r+1}\,y_2^{2(n-j_1)-r+1} \equiv 0.
\end{aligned} \tag{37}$$

The terms of the second line vanish by themselves. If we take

$$s_1^{(n+2)} = \sum_{k=0}^{n+2} e_k\, y_1^k\, y_2^{n+2-k}, \qquad s_2^{(m+2)} = \sum_{j=2}^{m+1} f_j\, y_1^j\, y_2^{m+2-j},$$

the identity (37) is easily reduced to

$$\sum_{j=0}^{n+2-2r} 2\,(j - j_1 + r - 1)\,(e_j - f_j)\, y_1^j\, y_2^{n+2-2r-j} \equiv -\,r\,(Q_1^{(1)} + Q_2^{(2)})\, y^{j_1-r+1}\, y_2^{n-j_1-r+1}.$$

For $j = j_1 - r + 1$ the coefficient of the left member vanishes, and the same must therefore be true of the right member. Therefore $Q_1^{(1)} + Q_2^{(2)} = 0$, since $r \neq 0$. For the other values of j we must have $e_j = f_j$. Hence

$$s_1^{(n+2)} = (y_1\, y_2)^r\, s_2^{(m+2)} + a y_1^{j_1+1}\, y_2^{n-j_1+1}, \tag{38}$$

where a is some constant.

Proceeding in this manner it can be shown by induction that

$$Q_1^{(i)} + Q_2^{(i)} = 0 \quad i = 1, \cdots, \infty, \qquad \text{and} \qquad S_1 = (y_1\, y_2)^r\, A\,(y_1\, y_2)\, S_2, \tag{39}$$

where A is an ordinary power series in the product $(y_1\, y_2)$, and this result is readily verified, for if we take

$$y_1' = y_1(+1 + Q^{(1)}\, y_1\, y_2 + Q^{(2)}\, y_1^2\, y_2^2 + \cdots),$$
$$y_2' = y_2(-1 - Q^{(1)}\, y_1\, y_2 - Q^{(2)}\, y_1^2\, y_2^2 - \cdots),$$
$$S_1 = AS_2,$$

where A is an arbitrary power series in $(y_1\, y_2)$, and substitute in (35) it is found at once that this equation is satisfied.

It follows therefore if there exists an integral of the original differential equations of the form

$$H = \frac{P_1(x_1\, x_2)}{P_2(x_1\, x_2)},$$

where P_1 and P_2 are power series in x_1 and x_2, that a linear transcendental substitution which transforms the differential equations into the form (31) is such as to make $Q_1^{(i)} + Q_2^{(i)} = 0$, $i = 1, \cdots, \infty$, and furthermore transforms the quotient P_1/P_2 into a power series in the product $(y_1\, y_2)$.

Returning now to equation (18) and putting

$$Q_1^{(1)} = -\,Q_2^{(1)} = Q, \qquad Q_1^{(2)} = -\,Q_2^{(2)} = 0,$$

it is clear that if $Q \neq 0$ the sum $(\alpha_1 + \beta_1)$ can be chosen so as to satisfy both equations. The difference $(\alpha_1 - \beta_1)$ remains arbitrary. From (18′) we have in general

$$Q_1^{(k+1)} = 0 = [(k-1)Q_1^{(1)} + kQ_2^{(1)}]\alpha_k - Q_1^{(1)}\beta_k + \text{known terms},$$

$$Q_2^{(k+1)} = 0 = -Q_2^{(1)}\alpha_k + [k(Q_1^{(1)} + (k-1)Q_2^{(1)}]\beta_k + \text{known terms}.$$

Since $Q_1^{(k+1)} + Q_2^{(k+1)} = 0$ the second equation is merely the negative of the first. Either of these equations gives

$$Q(\alpha_k + \beta_k) = \text{known terms}.$$

Thus $\alpha_k + \beta_k$ is determined by the choice $Q_1^{(k+1)} = Q_2^{(k+1)} = 0$. The difference $\alpha_k - \beta_k$ is undetermined. There are therefore infinitely many undetermined constants in the transformation, but whatever may be the value of these constants the reduced differential equations are

$$\begin{aligned} y_1' &= +y_1(1 + Qy_1y_2), \\ y_2' &= -y_2(1 + Qy_1y_2). \end{aligned} \tag{40}$$

From these equations are derived the integral, $y_1y_2 = c$, and therefore

$$\begin{aligned} y_1 &= \sqrt{c}\,e^{(1+Qc)(t-t_0)}, \\ y_2 &= \sqrt{c}\,e^{-(1+Qc)(t-t_0)}. \end{aligned} \tag{41}$$

Proof of Convergence. In the substitution (2) we will denote the sum of all terms in x_1 which have the form $y_1(y_1y_2)^s$ by $y_1\phi_1$, and the sum of all terms in x_2 which have the form $y_2(y_1y_2)^s$ by $y_2\phi_2$. Then we can write

$$\begin{aligned} x_1 &= y_1(1 + \phi_1) + f_1^*, \\ x_2 &= y_2(1 + \phi_2) + f_2^*; \end{aligned} \tag{42}$$

and from the process just described it is seen that the sum $\phi_1 + \phi_2$ is determined by the condition that the reduced equations shall have the form (40), while the difference $\phi_1 - \phi_2$ is undetermined.

The result of substituting (40) and (42) in the original equations (1) is

$$\begin{aligned} y_1\frac{\partial f_1^*}{\partial y_1} - y_2\frac{\partial f_1^*}{\partial y_2} - f_1^* &= \frac{X_1 + Qy_1^2y_2(1+\phi_1) + Qy_1y_2f_1^*}{1 - Qy_1y_2} = R_1, \\ y_1\frac{\partial f_2^*}{\partial y_1} - y_2\frac{\partial f_2^*}{\partial y_2} - f_2^* &= -\frac{X_2 + Qy_1y_2^2(1+\phi_2) + Qy_1y_2f_2^*}{1 - Qy_1y_2} = R_2, \end{aligned} \tag{43}$$

since

$$y_1\frac{\partial(y_1\phi_1)}{\partial y_1} - y_2\frac{\partial(y_1\phi_1)}{\partial y_2} - (y_1\phi_1) \equiv 0,$$

$$y_1\frac{\partial(y_2\phi_2)}{\partial y_1} - y_2\frac{\partial(y_2\phi_2)}{\partial y_2} + (y_2\phi_2) \equiv 0.$$

Taking now

$$f_1^* = \sum\sum m_{ij}\, y_1^i\, y_2^j, \qquad i+j \geqq 2 \quad \text{but} \quad i \neq j+1,$$
$$f_2^* = \sum\sum n_{ij}\, y_1^i\, y_2^j, \qquad i+j \geqq 2 \quad \text{but} \quad j \neq i+1,$$

there results

$$\sum\sum (i-j-1)\, m_{ij}\, y_1^i\, y_2^j = R_1, \tag{44}$$
$$\sum\sum (i-j+1)\, n_{ij}\, y_1^i\, y_2^j = R_2,$$

from which the coefficients of the left members can be computed successively if Q is properly chosen.

If, however, we add to the first equation

$$\sum m_{i,\,i-1}\, y_1^i\, y_2^{i-1} = y_1(1+\phi_1),$$

and subtract

$$\sum n_{i,\,i+1}\, y_1^i\, y_2^{i+1} = y_2(1+\phi_2),$$

from the second equation we have

$$\sum m_{i,\,i-1}\, y_1^i\, y_2^{i-1} + \sum'(i-j-1)\, m_{ij}\, y_1^i\, y_2^j = +\, y_1(1+\phi_1) + R_1, \tag{45}$$
$$\sum n_{i,\,i+1}\, y_1^i\, y_2^{i+1} + \sum'(i-j+1)\, n_{ij}\, y_1^i\, y_2^j = -\, y_2(1+\phi_2) - R_2.$$

The coefficients of the left members of these equations (45) also can be determined so that the equations are satisfied identically whatever ϕ_1 and ϕ_2 in the right members may be, and the solutions thus obtained are convergent; for the solutions thus obtained are dominated by the solutions of the equations which are obtained by replacing each $(i-j+1)$ in the left members of the first equation by $+1$ and each $(i-j+1)$ of the second equation by -1 and taking all the coefficients in the expansions of R_1 and R_2 with the positive sign, and the solutions of the equations thus modified are known to be convergent.

The coefficients m_{ij} and n_{ij} thus obtained from (45) will be functions of ϕ_1 and ϕ_2. But we must have

$$\sum m_{i,\,i-1}\, y_1^i\, y_2^{i-1} = y_1(1+\phi_1), \tag{46}$$
$$\sum n_{i,\,i+1}\, y_1^i\, y_2^{i+1} = y_2(1+\phi_2),$$

the left members of which are convergent series. It will be observed that in (45) the factor $(1+\phi_1)$ is always associated with y_1 except in the product $(y_1\, y_2)$ and the factor $(1+\phi_2)$ is associated in the same way with y_2. Hence the first equation of (46) contains $y_1(1+\phi_1)$ as a factor and the second contains $y_2(1+\phi_2)$ as a factor. Removing these factors and giving the m_{ij} and n_{ij} the values obtained from the solution of (45) we get

$$0 = [\,Q(1+\phi_1)(1+\phi_2) - Q\,]\, y_1\, y_2 + [\,]\, y_1^2\, y_2^2 + \cdots,$$
$$0 = [\,Q(1+\phi_1)(1+\phi_2) - Q\,]\, y_1\, y_2 + [\,]\, y_1^2\, y_2^2 + \cdots.$$

On removing the factor $y_1 y_2$ we have

$$0 = Q(\phi_1 + \phi_2) + [\,]\, y_1 y_2 + \cdots,$$
$$0 = Q(\phi_1 + \phi_2) + [\,]\, y_1 y_2 + \cdots. \tag{47}$$

These equations impose a condition only on $\phi_1 + \phi_2$ and they must therefore be identical. We have already seen that we can take $\phi_1 - \phi_2$ arbitrary. Hence we have

$$0 = Q(\phi_1 + \phi_2) + [\,]\, y_1 y_2 + \cdots,$$
$$0 = \phi_1 - \phi_2 + y_1 y_2 \psi(y_1 y_2) + \cdots. \tag{48}$$

If the arbitrary function ψ is given and is a convergent ordinary power series equations (48) admit a unique convergent solution for ϕ_1 and ϕ_2.

It might seem since m_{ij} and n_{ij} are functions of ϕ_1 and ϕ_2 that when the values thus derived for ϕ_1 and ϕ_2 are substituted in f_1^* and f_2^* they would then contain terms of the type $y_1^i y_2^{i-1}$, $y_1^i y_2^{i+1}$ respectively, but this is clearly not so, since the difference between the exponents of y_1 and y_2 in any term is not altered by multiplying the term by any power of $(y_1 y_2)$.

Example.—As an illustration of this method of solution let us consider the equation $d^2\phi/dt^2 + \sin\phi = 0$, whence $d^2\phi/dt^2 = -\phi + \frac{1}{6}\phi^3 - \frac{1}{120}\phi^5 + \cdots$. This equation will be reduced to the normal form if we take

$$\phi = \tfrac{1}{2}(\xi_1 + i\xi_2), \qquad d\phi/dt = \tfrac{1}{2}(i\xi_1 + \xi_2),$$

where $i = \sqrt{-1}$. The equations then are, if we take $it = \tau$,

$$\xi_1' = \xi_1 - f, \qquad \xi_2' = -\xi_2 - if,$$

where

$$f = \tfrac{1}{48}(\xi_1 + i\xi_2)^3 - \tfrac{1}{3840}(\xi_1 + i\xi_2)^5 + \cdots.$$

On making the linear-transcendental substitution

$$\xi_1 = \eta_1 + \left[-\frac{1}{96}\eta_1^3 - \frac{3i}{128}\eta_1^2\eta_2 - \frac{1}{32}\eta_1\eta_2^2 - \frac{i}{192}\eta_2^3 \right] + \cdots,$$
$$\xi_2 = \eta_2 + \left[-\frac{i}{192}\eta_1^3 + \frac{1}{32}\eta_1^2\eta_2 - \frac{3i}{128}\eta_1\eta_2^2 + \frac{1}{96}\eta_2^3 \right] + \cdots,$$

the reduced equations are found to be

$$\eta_1' = \eta_1\left(1 - \frac{i}{16}\eta_1\eta_2\right), \qquad \eta_2' = -\eta_2\left(1 - \frac{i}{16}\eta_1\eta_2\right).$$

The integral of these equations is $\eta_1\eta_2 = c$. It is convenient to take

$$c = -16i\gamma.$$

Then we find

$$\eta_1 = \sqrt{c}\, e^{(1-\gamma)\tau+\alpha}, \qquad \eta_2 = \sqrt{c}\, e^{-(1-\gamma)\tau-\alpha}.$$

If now we take $\alpha = -\pi i/4$ and replace τ by it we find

$$\phi = \tfrac{1}{2}(\xi_1 + i\xi_2) = [4 \sin (1-\gamma)t]\,\gamma^{1/2} + [\tfrac{1}{2} \sin (1-\gamma)t + \tfrac{1}{3} \sin 3(1-\gamma)t]\,\gamma^{3/2} + \cdots.$$

Since t does not enter explicitly in the original differential equation one can replace t by $(t - t_0)$ in the above solution in which $\phi = 0$ for $t = 0$.

INVARIANTS OF DIFFERENTIAL CONFIGURATIONS IN THE PLANE*

BY

EUGENE FRANCIS SIMONDS

INTRODUCTION

Suppose we have n points (x_ν, y_ν), $(\nu = 1, 2, 3, \cdots, n)$, in mutually general positions,† and let us assume that through each point (x_ν, y_ν) there pass λ_ν regular analytic elements of order r. The elements belonging to any point may have contact of zero order—that is, merely intersect—or some of them may have contact with each other of any order less than r. The totality of the elements considered will be called a *differential configuration of order r*.

If such a differential configuration be subjected to a group of point transformations, it will be changed always into another of the same type, that is, having the same constants n, λ_ν, and the order of contact of any two of the elements preserved. The numbers n, λ_ν are arithmetic invariants of the configuration under the group of point transformations, but will not necessarily be so under a group of contact transformations.

Rabut‡ appears to have been the first to pay systematic attention to the *differential* invariants of such configurations. He considered two groups—the group of all point transformations in the plane, and that of all contact transformations in the plane. He claimed for his method that it gave not only all configurations having invariants—both relative and absolute—but also all possible types of such invariants.

Doubt was thrown on the generality of this method by Kasner,§ who found (in another connection) a type of invariant which Rabut had declared impossible, and discussed a geometric interpretation of the same. (Rabut's actual error is pointed out by the writer in the present paper.)

* Presented to the Society, April 28, 1917.

† With regard to a group of transformations: the configuration of the points has no *special* property under the group.

‡ *Théorie des invariants universelles*, Journal de l'école polytechnique, ser. 2, vol. 4 (1898), pp. 137–203.

§ *The geometry of differential elements of the second order will respect to the group of all point transformations*, American Journal of Mathematics, vol. 28 (1906), p. 210.

In the article referred to Kasner also gave some invariants for the conformal group.*

The present paper deals only with the question of absolute invariants, more particularly with the types of configurations possessing them.

In Part I the problem is solved for any finite group of point transformations. It is shown that there are no essential invariants other than those of one or two elements. The invariants of the linear homogeneous group are found.

Part II treats of infinite groups of point transformations. In Section 1 it is shown that a single regular element has no invariant of order higher than zero. In Section 2 it is shown that every infinite group has invariants; but only for such configurations all of whose elements pass through a common point. The result admits of an exception, which is the subject of Section 3. In Section 4 we distinguish between groups involving arbitrary functions of only one variable, and those which involve at least one arbitrary function of two variables. Section 5 is devoted to a short consideration of the entire group of point transformations and of the area-preserving group. It is found that $2n + 2$ curves having contact of zero order have an invariant of order n under the entire group. For the area-preserving group the number is $n + 3$. Three types of invariants of lowest orders are given for each group. In Section 6 the equilong group is taken to illustrate Sections 3 and 4. We find that three curves have a complete series of invariants—one of each order—and obtain a differential operator by which they are all derived from the first.

Part III is concerned with contact transformations. The results obtained for point transformations have their counterpart for the contact transformations. The exceptional groups corresponding to those of Section 3, Part II, are found to be reducible to the latter by contact transformations. Rabut's general theorem about the entire group of contact transformations is verified by examination of the equivalent complete system of infinitesimal transformations.

In general, the method employed is that of infinitesimal transformations, following the Lie theory.

The writer has great pleasure in thanking Professor Kasner for kind encouragement and helpful criticism.

I. The invariants of differential configurations in the plane under finite groups of point transformations

We shall begin with a configuration of n points such that all the λ_ν elements through any point (x_ν, y_ν) have contact of zero order. Let the equation of

* Higher invariants of two elements for both the conformal and equilong groups are discussed by Kasner, *Conformal Geometry, Proceedings of the fifth international congress*, Cambridge (1912), vol. 2, pp. 81–87. See also other papers in these Transactions, vol. 16 (1915), pp. 333–349, and Bulletin of the American Mathematical Society, vol. 23 (1917), pp. 341–347.

any element through (x_ν, y_ν) be

$$y_{\nu,j} = y_{\nu,j}(x_\nu) \qquad (j = 1, 2, 3, \cdots, \lambda_\nu).$$

Consider an r-parameter group of point-transformations

$$X = \Phi(x, y, a_1, \cdots, a_r), \tag{1}$$
$$Y = \Psi(x, y, a_1, \cdots, a_r),$$

which does not degenerate in the neighborhood of the points (x_ν, y_ν). Let the infinitesimal transformations of the group be

$$X_\rho f \equiv \xi_\rho(x, y)\frac{\partial f}{\partial x} + \eta_\rho(x, y)\frac{\partial f}{\partial y} \qquad (\rho = 1, 2, \cdots, r).$$

Under an infinitesimal transformation $X_\rho f$ all the quantities

$$x_\nu, y_\nu, \cdots, y^{(k)}_{\nu,j}{}^{*}$$
$$(\nu = 1, 2, \cdots, n;\ j = 1, 2, \cdots, \lambda_\nu;\ k = 1, 2, \cdots, r)$$

are subjected to increments

$$\delta x_\nu = \xi_\rho(x_\nu, y_\nu)\,\delta t,$$
$$\delta y_\nu = \eta_\rho(x_\nu, y_\nu)\,\delta t,$$
$$\delta y'_{\nu,j} = \left\{\frac{\partial \eta_\rho(x_\nu, y_\nu)}{\partial x_\nu} + y'_{\nu,j}\left(\frac{\partial \eta_\rho}{\partial y_\nu} - \frac{\partial \xi_\rho}{\partial x_\nu}\right) - y'^{2}_{\nu,j}\frac{\partial \xi'_\rho}{\partial y_\nu}\right\}\delta t \equiv \eta'_{\rho,\nu,j}\,\delta t,$$
$$\cdots = \cdots$$
$$\delta y^{(k)}_{\nu,j} = \left\{\frac{d\eta^{(k-1)}_{\rho,\nu,j}}{dx_\nu} - y^{(k)}_{\nu,j}\frac{d\xi_\rho}{dx_\nu}\right\}\delta t \equiv \eta^{(k)}_{\rho,\nu,j}\,\delta t,$$
$$\cdots$$
$$(\nu = 1, 2, \cdots, n;\ j = 1, \cdots, \lambda_\nu;\ k = 2, 3, \cdots, r).$$

Any function $f(x_\nu, y_\nu, \cdots, y^{(k)}_{\nu j}, \cdots)$ is subjected to the increment $(X^{(r)}_\rho f)\,\delta t$, where

$$X^{(r)}_\rho f \equiv \sum_{\nu=1}^{n}\xi_\rho(x_\nu, y_\nu)\frac{\partial f}{\partial x_\nu} + \sum_{\nu=1}^{n}\eta_\rho\frac{\partial f}{\partial y_\nu} + \sum_{\nu=1}^{n}\sum_{j=1}^{\lambda_\nu}\sum_{k=1}^{r}\eta^{(k)}_{\rho,\nu,j}\frac{\partial f}{\partial y^{(k)}_{\nu,j}}. \tag{2}$$

The function $f(x_\nu, y_\nu, \cdots, y^{(k)}_{\nu,j}, \cdots)$ will be invariant under the group provided

$$X^{(r)}_\rho f \equiv 0 \qquad (\rho = 1, 2, \cdots, r).$$

Now by hypothesis the r independent infinitesimal transformations generate the group (1). They therefore satisfy two by two relations of the form

$$(X_i X_k) = \sum_{s=1}^{r} c_{iks} X_s f$$
$$(i, k = 1, 2, \cdots, r;\ c_{iks} = \text{constant}).$$

* We denote by $f^{(k)}$ the kth derivative of $f(x)$.

It is well known that the extended infinitesimal transformations $X_\rho^{(r)} f$ satisfy the analogous relations

$$(X_i^{(r)} X_k^{(r)}) = \sum_{s=1}^{r} c_{iks} X_s^{(r)} f$$

and generate an r-parameter group in the variables $x_\nu, y_\nu, \cdots, y_{\nu j}^k, \cdots$.

When equated to zero the expressions (2) form a complete system of r partial differential equations in the $2n + r\sum\lambda_\nu$ variables $x_\nu, y_\nu, \cdots, y_{\nu,j}^{(k)}, \cdots$. From the theory of complete systems we know that the number of independent solutions is

$$2n - r + r\sum_{\nu=1}^{n}\lambda_\nu.$$

These solutions are independent invariants of our differential configuration under the group of point transformations.

Of these invariants a certain number may be accounted for by taking them to be the ordinary differential invariants of single elements of the configuration. Each element has two of order not greater than $r - J_r$ and J_{r-1}; the former is of order r, the latter of order k ($< r$). If the group is transitive $k > 0$; if it is intransitive $k = 0$. The remaining independent invariants must then involve at least two elements. We shall refer to them as "mixed invariants." The number in the case of an r-parameter transitive group is

$$N_T = (2n - r) + (r - 2)\sum\lambda_\nu,$$

and in the case of an intransitive group

$$N_I = (n - r) + (r - 1)\sum\lambda_\nu.$$

We have so far supposed the elements through each of the n points to have contact of zero order. If this is not so let us assume the number of elements through the point (x_ν, y_ν) when the highest order considered is λ to be $n_{\nu,\lambda}$. The number of variables is now

$$2n + \sum_{\nu=1}^{n}\sum_{\lambda=1}^{r} n_{\nu,\lambda}.$$

The number of independent solutions is consequently

$$(2n - r) + \sum_{\nu=1}^{n}\sum_{\lambda=1}^{r} n_{\nu,\lambda}.$$

For a transitive group the number of ordinary differential invariants will be

$$\sum_{\nu=1}^{n}(n_{\nu,r} + n_{\nu,k}),$$

where k is the order of J_{r-1}; in this case, the number of mixed invariants is

$$N = (2n - r) - \sum_{\nu=1}^{n} (n_{\nu,r} + n_{\nu,k}) + \sum_{\nu=1}^{n} \sum_{\lambda=1}^{r} n_{\nu,\lambda}.$$

For an intransitive group we replace $\sum n_{\nu,k}$ by n.

The difference between this number N and the number of essential mixed invariants of the n "bunches" considered separately will be the number of independent essential invariants involving elements from more than one bunch. The latter is

$$n(2 - r) - \sum_{\nu=1}^{n} (n_{\nu,r} + n_{\nu,k}) + \sum_{\nu=1}^{n} \sum_{\lambda=1}^{r} n_{\nu,\lambda}.$$

The difference is therefore $(n - 1)r$, whether the group be transitive or not. Now two elements chosen from different bunches have r independent mixed invariants; hence the $(n - 1)r$ invariants may clearly be taken to be those of $n - 1$ configurations of two elements only, the elements of a pair being chosen from different bunches. It follows that the only essential invariants other than those of individual bunches may be taken to be those of two elements.

Let the first-order elements of a bunch through the point (x, y) be denoted by $y'_1, y'_2, \cdots, y'_{m_1}$, and let the number of elements of order r having the common slope y'_i be $n_i^{(0)}$. The configuration consisting of these elements alone has, say, $N_i^{(0)}$ invariants, whereas the whole bunch has N. We have, then,

$$\sum N_i^{(0)} = V - m_1 r + 2(m_1 - 1),$$

$$N = V - r,$$

where V is the number of variables in the whole bunch. Subtracting, we have

$$N - \sum N_i^{(0)} = (m_1 - 1)(r - 2). \tag{3}$$

If the group is transitive, these $(m_1 - 1)(r - 2)$ invariants are the independent invariants of any set of m_1 elements of order r, and may be taken to be the $m_1 - 1$ sets of $(r - 2)$ independent invariants of configurations consisting of pairs of elements having contact of order zero. If the group is intransitive the equation (3) takes the form

$$N - (\sum N_i^{(0)} - \overline{m - 1}) = (m_1 - 1)(r - 1),$$

showing that there are $(m_1 - 1)(r - 1)$ invariants of the above type. All other invariants of the bunch can be taken to be invariants of one or other of the sets $N_i^{(0)}$, that is, of configurations all of whose elements have contact of at least the first order.

Continuing the method of exhaustion for bunches all of whose elements have contact of the first order, we find in the same way that there are no essential

invariants of elements having contact of the first order other than those of *pairs* of elements.

And in general, if $y_1^{(s)}, y_2^{(s)}, \cdots, y_{m_s}^{(s)}$ be the only distinct elements of order s of the bunch, and if $N_i^{(s)}$ be the number of independent invariants of the configuration having the common element $y_i^{(s)}$ we have

$$N - \sum N_i^{(s)} = (m_s - 1)(r - s - 2).$$

In the case of a group where the order k of J_{r-1} is greater than s, the number $(m_s - 1)(r - s - 2)$ is the number of independent invariants of m_s elements of order r having contact of order s, exclusive of the types J_r and J_{r-1}. These invariants may be taken to be those of pairs of elements. If $k \leqq s$, instead of $\sum N_i^{(s)}$ we must subtract $\sum N_i^{(s)} - \overline{m_s - 1}$ from N, and we get in that case $(m_s - 1)(r - s - 1)$ invariants of *pairs* of elements having contact of order s.

We have, then, the following result for elements of order $\geqslant r$:

Under an r-parameter group of point transformations in the plane there are no essential invariants of order less than or equal to r other than those belonging to configurations of one or two elements.

So far we have spoken of elements of order $\geqslant r$. If we consider all elements of order $r + r'$, the number of additional independent invariants of any of the configurations considered can be accounted for by taking them to be the invariants of single elements, of which each curve has r' in addition to J_r and J_{r-1}. They are

$$J_{r+1} = \frac{dJ_r}{dJ_{r-1}}, \qquad J_{r+2} = \frac{dJ_{r+1}}{dJ_r}, \qquad \cdots, \qquad J_{r+r'} = \frac{dJ_{r+r'-1}}{dJ_{r+r'-2}}.$$

The final statement for all configurations is as follows:

An r-parameter group of point transformations in the plane has no essential mixed invariants other than those of two elements.

It will be convenient to distinguish symbolically the different types of configurations of two elements. We represent by $(1:1)_r$ that consisting of two distinct points (x_1, y_1) and (x_2, y_2) with an element of order r through each. The configuration $(2)_r^s$ is that one consisting of two elements of order r through a single point (x, y), the elements having contact of order s. If the elements have contact of zero order the symbol is $(2)_r^0$. A single element of order r is represented by $(1)_r$.

We may illustrate the foregoing by reference to the linear homogeneous group

$$X = a_1 x + a_2 y, \qquad Y = a_3 x + a_4 y.$$

The four mixed invariants of $(1:1)_4$ are

$$I_1 = \frac{y_2 - x_2 y_1'}{y_1 - x_1 y_1'}; \qquad I_2 = \frac{y_2 - x_2 y_2'}{y_1 - x_1 y_2'};$$

$$I_3 = \frac{y_1''(x_1 y_2 - x_2 y_1)^2}{[(y_2 - y_1) - (x_2 - x_1) y_1']^3}; \qquad I_4 = \frac{y_2''(x_1 y_2 - x_2 y_1)^2}{[(y_2 - y_1) - (x_2 - x_1) y_2']^3}.$$

We can obtain the two mixed invariants of $(2)_4^0$ from the above by seeking a function $f(I_1, I_2, I_3, I_4)$ which shall be independent of x_2, y_2. Such a function, if it exists, will be a solution of the pair of equations

$$\sum \frac{\partial f}{\partial I_i} \frac{\partial I_i}{\partial x_2} = 0, \qquad \sum \frac{\partial f}{\partial I_i} \frac{\partial I_i}{\partial y_2} = 0.$$

Two independent solutions are

$$I_1^{(0)} = \frac{(I_1 - I_2)^2}{I_3 (I_1 - 1)^3} = \frac{y_1''(y_1 - x_1 y_2')^2}{(y_1' - y_2')^2 (y_1 - x_1 y_1')},$$

$$I_2^{(0)} = \frac{(I_1 - I_2)^2}{I_4 (I_2 - 1)^3} = \frac{y_2''(y_1 - x_1 y_1')^2}{(y_1' - y_2')^2 (y_1 - x_1 y_2')}.$$

Finally, to obtain the unique mixed invariant of $(2)_4^0$, we divide $I_1^{(0)}$ by $I_2^{(0)}$ to eliminate the factor $y_1' - y_2'$, and put $y_2' = y_1'$. The result is

$$I_1^{(1)} = \frac{y_1''}{y_2''}.$$

The types I_1, I_2, I_3, I_4; $I_1^{(0)}, I_2^{(0)}$; $I_1^{(1)}$ are the only essential ones belonging to the linear homogeneous group apart from the series of invariants of a single element.

II. Infinite groups of point transformations

We consider the same problem when our group of transformations, viz.,

$$X = \Phi(x, y), \qquad Y = \Psi(x, y), \tag{1}$$

is infinite. The infinitesimal transformations of such a group will be of the form

$$\xi(x, y) \frac{\partial f}{\partial x} + \eta(x, y) \frac{\partial f}{\partial y},$$

where ξ and η are solutions of a system of linear partial differential equations*

$$A_i \xi + B_i \eta + C_i \xi_x + D_i \xi_y + E_i \eta_x + F_i \eta_y + \cdots = 0. \tag{2}$$

The general solution of this system involves one or more arbitrary functions either of one variable or of two. The most general case is that of the whole

* Lie, Mathematische Annalen, vol. 24 (1884), p. 553.

group of point transformations; the system (2) reduces to the identity $0 = 0$ and both ξ and η are arbitrary functions of x and y.

If it be required to find the differential invariants of the group (1), say of order n, we begin by "extending" the infinitesimal transformations, i. e., we calculate the increments of y', y'', $\cdots$, $y^{(n)}$. Let us call them, as is usual, $\eta'\,\delta t$, $\eta''\,\delta t$, $\cdots$, $\eta^{(n)}\,\delta t$. We give below such terms as are important in what is to follow.

$$
\begin{aligned}
&\eta' = \eta_x + (\eta_y - \xi_x)\,y' - \xi_y\,y'^2,\\
&\eta'' = \eta_{xx} + (2\eta_{xy} - \xi_{xx})\,y' + (\eta_{yy} - 2\xi_{xy})\,y'^2 - \xi_{yy}\,y'^3\\
&\qquad\qquad + (\eta_y - 2\xi_x)\,y'' - 3\xi_y\,y'\,y'',\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,\\
&\eta^{(n)} = \left\{ \eta_{n,0} + \sum_{k=1}^{n} (c_k\,\eta_{n-k,k} - c_{k-1}\,\xi_{n-k+1,k-1})\,y'^k - \xi_{0,n}\,y'^{n+1} \right\}\\
&\qquad + \left\{ \sum_{k=1}^{n-1} k\,(\bar{c}_k\,\eta_{n-k-1,k} - \bar{c}_{k-1}\,\xi_{n-k,k-1})\,y'^{k-1}\,y'' - n\xi_{0,n-1}\,y'^{n-1}\,y'' \right\}\\
&\qquad + \{(\eta_{n-2,1} - 2\xi_{n-1,0})\,y'' + (\overline{n-2}\,\eta_{n-3,2}\\
&\qquad - \overline{2n-1}\,\xi_{n-2,1})\,y'\,y'' + \cdots\\
&\qquad + (2\eta_{0,n-1} - 3\overline{n-1}\,\xi_{1\,n-2})\,y'^{n-2}\,y'' - 3\xi_{0,n-1}\,y'^{n-1}\,y''\} + \cdots.
\end{aligned}
\tag{3}
$$

Here $f_{n-k,k}$ is the derivative of the nth order obtained by differentiating $f(x, y)$ $n-k$ times with respect to x, and k times with respect to y; c_k is the coefficient of x^k in the expansion of $(1-x)^n$; and $\bar{c}_k$ is the coefficient of x^k in the expansion of $(1-x)^{n-1}$. In the expression for $\eta^{(n)}$, three sets of terms are given. The first is a polynomial in y', whose coefficients involve all the partial derivatives of ξ and η of order n. The coefficients in the second already occur in $\eta^{(n-1)}$; the third involves $n-1$ functions of the derivatives of order $n-1$ which are all independent of those already occurring in $\eta^{(n-1)}$. The three sets of terms contain all the terms of $\eta^{(n)}$ whose coefficients involve derivatives of ξ and η of orders $n-1$ and n.

Any function $f(x, y, \cdots, y^{(n)})$ is subjected by the infinitesimal transformation ξ, η to an increment

$$
\left\{ \xi\frac{\partial f}{\partial x} + \eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'} + \cdots + \eta^{(n)}\frac{\partial f}{\partial y^{(n)}} \right\} \delta t,
$$

where ξ and η are subject to the conditions imposed by the defining system (2). Now the equations of this system enable us to express certain of the functions ξ, η, ξ_x, ξ_y, η_x, η_y, $\cdots$ in terms of the rest. Making the resulting substitutions we reduce the number of independent functions occurring in the ex-

pressions $\xi, \eta, \eta', \eta'', \cdots, \eta^{(n)}, \cdots$. The system (3) is then equivalent to a complete system of, say, $N^{(n)}$ infinitesimal transformations

$$\alpha_j \frac{\partial f}{\partial x} + \beta_j \frac{\partial f}{\partial y} + \cdots + \beta_j^{(n)} \frac{\partial f}{\partial y^{(n)}} \quad (j = 1, 2, \cdots, N^{(n)}), \tag{4}$$

where α_j, β_j are functions of x and y, and $\beta_j^{(n)}$ is a function of $x, y, y', \cdots, y^{(n)}$. If an element of order n has invariants under the group, they will be solutions of the equations obtained by setting the expressions (4) equal to zero.

1. **A single regular element.** Considering first a single element, we know that the system sometimes has a solution of order zero. For example if $\alpha_j = 0$, a solution will be x. The largest group with this invariant is

$$X = x, \qquad Y = \Psi(x, y),$$

Ψ being an arbitrary function.

We are further able to assert the following:

No infinite group of point transformations can have an invariant of a single element of order greater than zero.

This result is implicit in the following theorem of Lie's:* "If an ordinary differential invariant of the rth order 'fixes' a ρ-parameter group of transformations, the number ρ has a finite upper limit."

For, by equating to a constant a differential invariant of order n, we get a differential equation of order n which is invariant. Thus, there is no infinite group with an invariant of order > 1. Moreover, if the group have an invariant $f(x, y, y')$, it will leave invariant the differential equation

$$f_x + f_y y' + f_{y'} y'' = 0,$$

and that in a special way, for it will leave invariant each of the ∞^1 families $f(x, y, y') = c$ of the ∞^2 integral-curves. Thus Lie's theorem excludes invariants of order 1.

The following is an independent proof, more in line with the other considerations of this chapter. Let

$$X = \Phi(x, y), \qquad Y = \Psi(x, y)$$

be a group of point transformations, and let

$$Uf \equiv \xi \frac{\partial f}{\partial x} + \eta \frac{\partial f}{\partial y}$$

be an infinitesimal transformation of the group. ξ and η are solutions of a system of defining-equations. Suppose the group to have a differential invariant of order $n\,(>0)$ e. g.,

$$\Omega(x, y, y', \cdots, y^{(n)}).$$

* Lie-Scheffers, *Continuierliche Gruppen*, Kap. 12, § 1.

Then it is necessary and sufficient that

$$U^{(n)}\,\Omega = 0,$$

where $U^{(n)}f$ is the n-fold extension of Uf. $U^{(n)}f$ is, in fact,

$$\xi\frac{\partial f}{\partial x} + \eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'} + \cdots + \eta^{(n)}\frac{\partial f}{\partial y^{(n)}},$$

where the expressions for $\eta', \eta'', \cdots, \eta^{(n)}$ are those given by equations (3).

We must have, then, for all values of $y', y'', \cdots, y^{(n)}$,

$$U^{(n)}\,\Omega = 0,$$

in virtue of the relations between ξ and η and their partial derivatives given by the defining-equations of the group.

Conversely, the defining-equations of the largest group having the invariant Ω are obtained from the identity

$$U^{(n)}\,\Omega = 0 \tag{5}$$

in the variables $y', y'', \cdots, y^{(n)}$, by equating to zero the coefficients of the different powers of $y', y'', \cdots, y^{(n)}$ and their products.

Solving the equation (5) for $\eta^{(n)}$ we have

$$\eta^{(n)} = -\left(\xi\frac{\partial\Omega}{\partial x} + \eta\frac{\partial\Omega}{\partial y} + \cdots + \eta^{(n-1)}\frac{\partial\Omega}{\partial y^{(n-1)}}\right) \div \frac{\partial\Omega}{\partial y^{(n)}}.$$

This is an identity in $y', y'', \cdots, y^{(n)}$. Let us now choose such a set of values for $y'', \cdots, y^{(n)}$ that $\partial\Omega/\partial y^{(n)}$ does not vanish. After substitution we have, making use of the expression for $\eta^{(n)}$ in equations (3),

$$\eta_{n,0} + (c_1\,\eta_{n-1,1} - \xi_{n,0})\,y' + \sum_{k=2}^{n}(c_k\,\eta_{n-k,k} - c_{k-1}\,\xi_{n-k+1,k-1})\,y'^k + \cdots + c_n\,\xi_{0,n}\,y'^{n+1} = \lambda(x, y, y'),$$

where λ involves partial derivatives of ξ and η of order lower than n. This is an identity in y'. If we choose $n+2$ different values of y', no one of which causes the denominator of λ to vanish, we obtain a system of $n+2$ linear independent equations in the $n+2$ quantities

$$\eta_{n,0},\ c_k\,\eta_{n-k,k} - c_{k-1}\,\xi_{n-k+1,k-1},\ \xi_{0,n}.$$

Solving, we get $n+2$ equations of the type

$$\begin{aligned} \eta_{n,0} &= \sigma_0, \\ c_k\,\eta_{n-k,k} - c_{k-1}\,\xi_{n-k+1,k-1} &= \sigma_k \qquad (k = 1, 2, \cdots, n), \\ \xi_{0,n} &= \sigma_{n+1}, \end{aligned} \tag{6}$$

where the σ's are functions of the partial derivatives of ξ and η of order lower than n, and also perhaps involve x and y explicitly.

If we differentiate these equations partially for x and y we get a system of $2(n+2)$ linear equations

$$
\begin{aligned}
(\eta_{n,\,0})_x &= (\sigma_0)_x, \\
(\eta_{n,\,0})_y &= (\sigma_0)_y, \\
(c_k\,\eta_{n-k,\,k} - c_{k-1}\,\xi_{n-k+1,\,k-1})_x &= (\sigma_k)_x, \\
(c_k\,\eta_{n-k,\,k} - c_{k-1}\,\xi_{n-k+1,\,k-1})_y &= (\sigma_k)_y, \\
(\xi_{0,\,n})_x &= (\sigma_{n+1})_x, \\
(\xi_{0,\,n})_y &= (\sigma_{n+1})_y,
\end{aligned}
\qquad (k = 1, 2, \cdots, n), \tag{7}
$$

in the $2(n+2)$ partial derivatives of ξ and η of order $n+1$, whose determinant has the value

$$(-1)^n (c_1 - c_1^2)(c_1 c_3 - c_2^2) \cdots (c_{n+2} c_n - c_{n+1}^2)$$

and hence does not vanish. The system consisting of (6) and (7) together has the following properties:

(*a*) The highest order of the partial derivatives of ξ and η occurring in the system being $n+1$, it is possible to express all the derivatives of order $n+1$ in terms of those of lower orders, and this is not the case for all the derivatives of any order less than $n+1$.

(*b*) It is not possible through partial differentiation of the system and combination of the resulting equations with each other and with those of the system to obtain any further equations expressing relations between the derivatives from the first to the nth order, independent of those already occurring in the system.

Now any system of partial differential equations possessing these two properties is integrable, and its general solution involves only a finite number of arbitrary constants.*

The whole system of defining-equations—of which the system (6) is but a part—will therefore have a solution-pair ξ, η, involving only a finite number of parameters.

The group with the invariant Ω is therefore finite.

2. **Configurations which have invariants.** Turning now to the differential configuration consisting of r distinct points (x_ν, y_ν), the number of elements in the "bunch" (if we may adopt the word) through (x_ν, y_ν) being λ_ν, we see at once a fundamental distinction between finite and infinite groups. The set of essential arbitrary functions ξ, η, ξ_x, η_x, ξ_y, η_y, $\cdots$ to be elimi-

* Lie, *Theorie der Transformationsgruppen*, vol. 1, Kap. 10.

nated occurs in each bunch with different arguments. Each bunch, therefore, has its own complete system, and the complete system belonging to the whole configuration is merely the combination of the complete systems of the individual bunches. Since the same is obviously true of the number of variables in the whole configuration, we have the following

THEOREM. *The only essential invariants, under an infinite group of point transformations, of a differential configuration consisting of several bunches of elements whose base-points are in mutually general positions are those of the individual bunches.*

We have to consider, then, the configuration $(\lambda)_n^0$ consisting of λ elements of order n having the point (x, y) in common. The equivalent complete system is

$$\alpha_j(x, y)\frac{\partial f}{\partial x} + \beta_j(x, y)\frac{\partial f}{\partial y} + \sum_{k=1}^{n} \beta_{j\mu}^{(k)} \frac{\partial f}{\partial y_\mu^{(k)}}$$
$$(j = 1, 2, \cdots, N_n;\ \mu = 1, 2, \cdots, \lambda).$$

For $\lambda = 1$, there is no invariant of order greater than zero, but since the number of equations in the complete system depends on the group, while λ is at our disposal, *it is clearly possible by sufficiently increasing* λ *to obtain a configuration which shall have an invariant of order* n.

The theorem that all groups of point-transformations have invariants of certain configurations is a result due to Rabut,* who also showed that the same was true for all groups of contact transformations.

3. **Certain imprimitive groups.** We have seen that an infinite group of point transformations has no essential invariants apart from those of individual bunches, provided the base-points of the bunches are in mutually general positions. In the case, however, of certain imprimitive groups if we choose the base-points (x_ν, y_ν) of the bunches on the same integral-curve of the invariant differential equation of the first order, it happens that there are essential invariants of configurations consisting of two or more bunches. All imprimitive groups of the plane are equivalent under the group of point transformations to the group

$$X = \Phi(x), \qquad Y = \Psi(x,).$$

We may ask for the most general form of the function Ψ so that the group may have invariants of the above type.

It is clear that the most general configuration of the desired type is that consisting of, say, r distinct points (x, y_ν) each with a single element passing through it. Let the infinitesimal transformations of the group be

$$\delta x = \xi(x)\,\delta t; \qquad \delta y = \eta(x, y)\,\delta t.$$

* Loc. cit.

Now η can involve no arbitrary function of y, which varies from element to element, since this would necessarily lead to the equations

$$\frac{\partial f}{\partial y_\nu} = 0, \qquad \frac{\partial f}{\partial y_\nu'} = 0, \qquad \cdots, \qquad \frac{\partial f}{\partial y_\nu^{(k)}} = 0;$$

η then, must be a linear expression of the form

$$\alpha_1(y)\beta_1(x) + \alpha_2(y)\beta_2(x) + \cdots + \alpha_k(y)\beta_k(x),$$

where the α's are definite functions of y, but the β's may be arbitrary functions of x. The equivalent complete system of the unextended group is then

$$\frac{\partial f}{\partial x}, \ \alpha_k \frac{\partial f}{\partial y} \qquad (k = 1, 2, \cdots).$$

Since these generate a finite group, the transformations

$$\alpha_k \frac{\partial f}{\partial y} \qquad (k = 1, 2, \cdots),$$

themselves generate a finite group in the variable y. There are, however, only three independent infinitesimal transformations in one variable generating a finite group viz.,

$$\frac{\partial f}{\partial y}, \ y\frac{\partial f}{\partial y}, \ y^2\frac{\partial f}{\partial y},$$

so that η is of the form

$$\beta_1(x) + y\beta_2(x) + y^2\beta_3(x).$$

Then the finite equations of the group are

$$X = \Phi(x), \qquad Y = \frac{\Psi_1(x) + \Psi_2(x)y}{\Psi_3(x) + \Psi_4(x)y},$$

Φ, Ψ_1, Ψ_2, Ψ_3, Ψ_4 being arbitrary functions of x, only three of the last four, of course, being essential.

This group has an invariant of four points having the same abscissa, viz.,

$$\frac{(y_1 - y_2)(y_3 - y_4)}{(y_2 - y_3)(y_4 - y_1)}.$$

4. **Number of elements having invariants.** There is also a distinction between groups whose equations involve arbitrary functions of two variables and those involving only arbitrary functions of one variable. In the former case, the number of "new" arbitrary functions after each successive differentiation is an increasing function of n. If λ_n be the lowest number of curves (having contact of order zero) with an invariant of order n, we have in succession (expressing algebraically the fact that the number of independent

solutions of a complete system is equal to the number of variables minus the number of independent equations)

$$\begin{aligned}
\lambda_1 &= N_1 + 1, \\
\lambda_2 + \lambda_1 &= N_2 + 2, \\
\lambda_3 + \lambda_2 + \lambda_1 &= N_3 + 3, \\
&\cdots\cdots\cdots\cdots \\
\lambda_n + \lambda_{n-1} + \cdots + \lambda_1 &= N_n + n.
\end{aligned}$$

Subtracting the last two we get

$$\lambda_n = 1 + (N_n - N_{n-1}), \tag{8}$$

which shows that λ_n is an increasing function of n. In the second case, $N_n - N_{n-1}$ is constant, and equal to the number of unrelated arbitrary functions in the unextended form of the group. From the category of functions which are " unrelated " to a given set of functions, we must exclude not only dependent functions, but those dependent on the *given set and their derivatives.* If the number of unrelated functions in the group is k, $k + 1$ curves have an invariant of each order after the lowest.

It was assumed in the foregoing that the curves of the configuration had contact of the lowest possible order, that is to say, that in the case of the groups considered in Section 3 all the elements pass through points which have but one coördinate in common, and in all other cases have contact of zero order. We may say then:

THEOREM. *If the elements of our configuration have contact of the lowest possible order consistent with the existence of differential invariants, and if* λ_n *be the lowest number of elements having an invariant of order n under an infinite group, two cases occur:*

(*a*) λ_n *is an increasing function of n if the group involves at least one arbitrary function of two variables;*

(*b*) λ_n *attains a final value* $k + 1$ *if the equations of the group involve k arbitrary functions all of one variable.*

5. **The group of all point-transformations and the group which preserves area.** As an example illustrating Section 2 and the first case of Section 4, we take the area-preserving group; to illustrate Section 3 and the second case of Section 4 we discuss the equilong group. First, however, we obtain a theorem about the entire group of point transformations.

The theory of this group was partly worked out by Rabut,* who obtained a set of results which he supposed to be perfectly general and all-inclusive, but which were shown to be not so by Kasner.† Among other correct results

* Loc. cit.

† Loc. cit.

Rabut shows (*a*) the invariants are free from x and y; (*b*) each bunch must be considered separately; (*c*) the invariants are of degree and of weight zero in all of the derivatives; (*d*) derivatives of the same order appear only in their differences two by two. He goes on, however, to say: "in order that $y_a^{(m)} - y_b^{(m)}$ may occur in an invariant, the curves a and b must have contact of order $m - 1$." The proof given* is fallacious, as the invariant obtained by Kasner showed. This error made an apparent simplification of the problem, whose complete solution, in fact, is not practicable, owing to the increasing complexity of the expressions for the successive derivatives of Y with regard to X. This (the chief obstacle to a complete solution of the problem) was entirely avoided by Rabut's erroneous assumption. His results, then, are far from being general, but apply only to a very limited class of configurations.

In equations (3) we found expressions for the increments of the variables $x, y, y', \cdots, y^{(n)}$ under the general infinitesimal transformation of the entire group, pointing out what arbitrary independent functions of x and y occurred in the various expressions. The number of these is $n(n+2)+2$. Thus the general infinitesimal transformation of the group is equivalent to a complete system of $n(n+2)+2$ infinitesimal transformations of which the first two are $\partial f/\partial x$, $\partial f/\partial y$. The formula (8) gives $\lambda_n = 2n+2$, and the complete system of equations for a bunch of $2n+2$ curves intersecting in a common point, is as follows:

$$\frac{\partial f}{\partial x} = 0; \qquad \frac{\partial f}{\partial y} = 0; \qquad \sum_{\iota=1}^{2n+1} \frac{\partial f}{\partial y_\iota'} = 0,$$

$$\sum_{\iota=1}^{2n+2} \left(y_\iota' \frac{\partial f}{\partial y_\iota'} + y_\iota'' \frac{\partial f}{\partial y_\iota''} + \cdots + y_\iota^{(n)} \frac{\partial f}{\partial y_\iota^{n}} \right) = 0,$$

$$\sum_{\iota=1}^{2n+2} \left(y_\iota' \frac{\partial f}{\partial y_\iota'} + 2y_\iota'' \frac{\partial f}{\partial y_\iota''} + \cdots + n y_\iota^{(n)} \frac{\partial f}{\partial y_\iota^{(n)}} \right) = 0,$$

$$\sum_{\iota=1}^{2n+2} \left(y_\iota'^2 \frac{\partial f}{\partial y_\iota'} + 3y_\iota' y_\iota'' \frac{\partial f}{\partial y_\iota''} + \cdots \right) = 0,$$

$$\cdots \cdots \cdots \cdots \cdots \cdots \cdots$$

$$\sum_{\iota=1}^{2n+2} y_\iota'^{\mu} \frac{\partial f}{\partial y_\iota^{(n)}} = 0 \qquad (\mu = 0, 1, 2, \cdots, n+1),$$

$$\sum_{\iota=1}^{2n+2} y_\iota'^{\nu} y_\iota'' \frac{\partial f}{\partial y_\iota^{(n)}} = 0 \qquad (\nu = 0, 1, 2, \cdots, n-2).$$

The first two of these equations show that x and y do not occur in an invariant, as already stated.

We must justify our use of the formula (8) by showing that the equations

* P. 166.

of the above system are independent. If we assume that this is the case for $n = k - I$, the matrix M_{k-1} of the system (which we may call S_{k-1}) has at least one determinant of order $k^2 + 1$ which does not vanish. Among the new rows and columns in S_k is found the matrix

$$\begin{vmatrix} 1, & 1, & 1, & \cdots, & 1 \\ y_1', & y_2', & y_3', & \cdots, & y_{2k+2}' \\ y_1'^2, & y_2'^2, & y_3'^2, & \cdots, & y_{2k+2}'^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ y_1'^{k+1}, & y_2'^{k+1}, & y_3'^{k+1}, & \cdots, & y_{2k+2}'^{k+1} \\ y_1'', & y_2'', & y_3'', & \cdots, & y_{2k+2}'' \\ y_1' y_1'', & y_2' y_2'', & y_3' y_3'', & \cdots, & y_{2k+2}' y_{2k+2}'' \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ y_1'^{k-2} y_1'', & y_2'^{k-2} y_2'', & y_3'^{k-2} y_3'', & \cdots, & y_{2k+2}'^{k-2} y_{2k+2}'^{k-2} \end{vmatrix}$$

no one of whose $(2k + 1)$-rowed determinants vanishes. Thus S_k has at least one $[k(k + 2) + 2]$-rowed determinant which does not vanish. If then, the system is independent for $n = k - 1$, it is independent also for $n = k$. It is known to be so for $n = 2$; hence the system is always independent. The number of independent solutions is n^2. Now four curves have an invariant of the first order, therefore the total number of invariants of the first order is $(2n + 2) - 3$. Those of the second order independent of each other and of those of the first order number $2n + 2 - 5$, and so on; adding together all the essential invariants of orders 1 to $n - 1$ we get $n^2 - 1$. The remaining invariant is of order n. We assert then the following

THEOREM. *The smallest number of curves having contact of zero order possessing an invariant of order n under the entire group of point transformations is* $2n + 2$.

This agrees with the known results for $n = 1$, $n = 2$, given by Kasner.*

The area-preserving group is defined by the equation

$$\xi_x + \eta_y = 0. \tag{9}$$

By differentiating this equation we obtain two equations of the second order, viz.,

$$\xi_{xx} + \eta_{xy} = 0, \qquad \xi_{xy} + \eta_{yy} = 0. \tag{10}$$

The relation (9) is without effect in reducing the number of equations in the complete system for the whole group of point transformations extended to the first order, for ξ_x and η_y occur only in the expression $\eta_y - \xi_x$. Similarly, the equations (10) have no effect in reducing the complete system extended

* Loc. cit.

to the second order. Since, however, ξ_x and η_y now occur in the combination $\eta_y - 2\xi_x$ which is independent of $\eta_y - \xi_x$, the equation (9) reduces the number of equations by 1. In general, the r equations of order r obtained from (9) by differentiation will not affect the complete system for the whole group extended to the rth order, but will reduce the system extended to the $(r+1)$th order by r equations. The complete system of the area-preserving group contains therefore $(1+2+3+\cdots+n-1)$ equations fewer than that of the entire group. The number of equations is therefore $\frac{1}{2}n(n+5)$. The formula (8) for λ_n then gives $n+3$. Hence

THEOREM. *The smallest number of curves having contact of zero order possessing an invariant of order n under the area-preserving group is* $n+3$.

The first type of invariant is obviously

$$I = \frac{(y_1' - y_2')(y_3' - y_4')}{(y_2' - y_3')(y_4' - y_1')},$$

which is the same as that of the entire group. We shall obtain the second by integrating the complete system. We have 7 equations in 10 variables:

$$U_1 f \equiv \sum \frac{\partial f}{\partial y_i'} = 0; \qquad U_2 f \equiv \sum \left(2y_i' \frac{\partial f}{\partial y_i'} + 3y_i'' \frac{\partial f}{\partial y_i''} \right) = 0;$$

$$U_3 f \equiv \sum \left(y_i'^2 \frac{\partial f}{\partial y_i'} + 3y_i' y_i'' \frac{\partial f}{\partial y_i''} \right) = 0; \qquad U_4 f \equiv \sum \frac{\partial f}{\partial y_i''} = 0;$$

$$U_5 f \equiv \sum y_i' \frac{\partial f}{\partial y_i''} = 0; \qquad U_6 f \equiv \sum y_i'^2 \frac{\partial f}{\partial y_i''} = 0; \qquad U_7 f \equiv \sum y_i'^3 \frac{\partial f}{\partial y_i''} = 0,$$

the summations extending from $i = 1$ to $i = 5$. By inspection we obtain five independent solutions of $U_1 = U_4 = U_5 = U_6 = U_7 = 0$, namely,

$$Z_j = y_1' - y_j' \quad (j = 2, 3, 4, 5), \qquad D_5 \equiv \begin{vmatrix} 1, & y_1', & y_1'^2, & y_1'^3, & y_1'' \\ 1, & y_2', & y_2'^2, & y_2'^3, & y_2'' \\ 1, & y_3', & y_3'^2, & y_3'^3, & y_3'' \\ 1, & y_4', & y_4'^2, & y_4'^3, & y_4'' \\ 1, & y_5', & y_5'^2, & y_5'^3, & y_5'' \end{vmatrix}.$$

Substituting in U_2 and U_3 we get

$$\sum_{j=2}^{5} 2Z_j \frac{\partial f}{\partial Z_j} + 15D_5 \frac{\partial f}{\partial D_5} = 0, \tag{11}$$

$$\sum_{j=2}^{5} (y_1' + y_j') Z_j \frac{\partial f}{\partial Z_j} + 3 \sum_{i=1}^{5} y_i' D_5 \frac{\partial f}{\partial D_5} = 0. \tag{12}$$

Multiplying equation (11) by y_1' and subtracting from (12)

$$\sum_{j=2}^{5} Z_j \frac{\partial f}{\partial Z_j} + 3\sum_{j=2}^{5} Z_j D_5 \frac{\partial f}{\partial D_5} = 0. \tag{13}$$

We obtain four independent solutions of (13), viz.,

$$\bar{Z}_k \equiv \frac{1}{Z_2} - \frac{1}{Z_k} (k = 3, 4, 5), \qquad \bar{D} = \frac{D_5^{1/3}}{Z_2 Z_3 Z_4 Z_5}.$$

Substituting in (11) we have

$$2\sum_{k=3}^{5} Z_k \frac{\partial f}{\partial \bar{Z}_k} + 3\bar{D}\frac{\partial f}{\partial \bar{D}} = 0. \tag{14}$$

Finally, three independent solutions of (14) are

$$\frac{\bar{Z}_3}{\bar{Z}_5}, \frac{\bar{Z}_4}{\bar{Z}_5}, \frac{\bar{D}^{1/3}}{\bar{Z}_3^{1/2}}.$$

The first two are of the type I; when raised to the 9th power and combined with the other two the third reduces to the form

$$J_5 = \frac{D_5^{2/3}(y_1' - y_2')(y_1' - y_3')}{(y_1' - y_4')^2 (y_1' - y_5')^2 (y_2' - y_3')^3}.$$

Reverting for a moment to the complete system for the entire group, we notice that it consists of 8 equations—6 being identical with those of the area-preserving group, but the other two

$$V_1 f \equiv \sum y_i' \frac{\partial f}{\partial y_i} = 0,$$

$$V_2 f \equiv \sum y_i'' \frac{\partial f}{\partial y_i''} = 0$$

being replaced in the area-preserving group by the single equation

$$\bar{U}_2 f \equiv (2V_1 + 3V_2) f = 0,$$

the summations now extending from 1 to 6, a change which we will indicate by the bar over the symbol Uf. Now any solution of $\bar{U}_2 f = 0$ which is homogeneous and of degree zero both in the y''s and the y'''s will be a solution of both $V_1 f = 0$ and $V_2 f = 0$. Now, we know two solutions of $\bar{U}_2 f = 0$, viz., J_5 and J_6, where J_6 is obtained from J_5 by the substitution of the subscript 6 for the subscript 5. These also satisfy the remaining equations of the complete system. A function of J_5 and J_6 fulfilling all the required con-

ditions is

$$\left(\frac{J_5}{J_6}\right)^3 \equiv \frac{(y_1' - y_6')^3 D_5}{(y_1' - y_5')^3 D_6},$$

D_6 being obtained from D_5 by the substitution of the subscript 6 for the subscript 5. This is the type of invariant obtained by Kasner*—though in a different form—which showed the inadequacy of Rabut's treatment of the entire group of point transformations.

The invariants of the entire group are always homogeneous of degree zero in all of the variables (though not necessarily in the variables of each order). This appears from the fourth equation of the complete system. The fifth equation shows that the weight of every invariant is zero, a result stated by Rabut. In the area-preserving group these two equations are replaced by the single equation

$$\sum 2y_\iota' \frac{\partial f}{\partial y_\iota'} + 3y_\iota'' \frac{\partial f}{\partial y_\iota''} + \cdots + (n+1) y_\iota^{(n)} \frac{\partial f}{\partial y_\iota^{(n)}} = 0,$$

which has a set of fundamental solutions

$$y_i'^{1/2} \div (y_j^{(k)})^{1/(k+1)},$$

showing that the following theorem is true.

THEOREM. *If the order of derivation of every variable occurring in an invariant of the area-preserving group be increased by unity, the result is an expression of weight zero.*

(After the above operation, both numerator and denominator of J_5 are of weight 14.)

The factors $y_\iota' - y_j'$ and D_5 occurring in the expressions J_5 and $J_5 \div J_6$ are relative invariants of the entire group. The magnifying factor of the first is

$$\frac{\Phi_x \Psi_y - \Psi_x \Phi_y}{(\Phi_x + \Phi_y y_\iota')(\Phi_x + \Phi_y y_j')},$$

while that of D_5 is

$$\frac{(\Phi_x \Psi_y - \Psi_x \Phi_y)^7}{\prod_{\iota=1}^{\iota=5} (\Phi_x + \Phi_y y_\iota')^3}.$$

In the case of the area-preserving group $\Phi_x \Psi_y - \Psi_x \Phi_y = 1$.

We give a third invariant of this latter group, which was obtained by the method described in the next section. A convenient form is

* P. 210.

$$K_6 = \frac{\begin{bmatrix}1\,2\,3\\1\,2\,3\end{bmatrix}^2}{\begin{bmatrix}1\,2\\1\,2\end{bmatrix}^4\begin{bmatrix}1\,3\\1\,2\end{bmatrix}^4} \cdot \frac{\begin{bmatrix}1\,2\,3\,4\,5\,6\\1\,2\,3\,4\,5\,9\end{bmatrix}\begin{bmatrix}1\,2\,3\,4\\1\,2\,3\,4\end{bmatrix} - \begin{bmatrix}1\,2\,3\,4\,5\,6\\1\,2\,3\,4\,5\,6\end{bmatrix}\begin{bmatrix}1\,2\,3\,4\\1\,6\,3\,4\end{bmatrix} - 2\begin{bmatrix}1\,2\,3\,4\,5\,6\\1\,2\,3\,4\,5\,7\end{bmatrix}\begin{bmatrix}1\,2\,3\,4\\1\,2\,6\,4\end{bmatrix} - 3\begin{bmatrix}1\,2\,3\,4\,5\,6\\1\,2\,3\,4\,5\,6\end{bmatrix}\begin{bmatrix}1\,2\,3\,4\\1\,2\,3\,6\end{bmatrix}}{\begin{bmatrix}1\,2\,3\,4\\1\,2\,3\,4\end{bmatrix}\begin{bmatrix}2\,3\,4\,5\,6\\1\,2\,3\,4\,5\end{bmatrix}},$$

where the quantities in brackets are determinants formed from the matrix whose ith row is the following:

$$1 \quad y_i' \quad y_i'^2 \quad y_i'^3 \quad y_i'^4 \quad y_i'' \quad y_i' y_i'' \quad y_i'^2 y_i'' \quad y_i''' ,$$

the upper and lower rows of numerals within each bracket indicating respectively the rows and columns of the matrix from which the determinant is constructed.

From this result we can build up a third-order invariant of the entire group for 8 curves. The complete system equivalent to the entire group contains the equations

$$W_1 f = \sum \left(y_i' \frac{\partial f}{\partial y_i'} + y_i'' \frac{\partial f}{\partial y_i''} + y_i''' \frac{\partial f}{\partial y_i'''} \right) = 0,$$

$$W_2 f = \sum y_i'' \frac{\partial f}{\partial y_i'''} = 0,$$

$$W_3 f = \sum y_i' y_i'' \frac{\partial f}{\partial y_i'''} = 0,$$

in addition to the twelve whose solutions we have found.

The configuration has three independent solutions of the third order under the area-preserving group, viz.,

$$K_6, \; K_7, \; K_8$$

where K_7, K_8 are obtained from K_6 by the substitution of the subscripts 7 and 8 respectively for the subscript 6. Let us seek a linear function of K_6, K_7, K_8 which shall satisfy the equations $W_2 f = 0$ and $W_3 f = 0$. If

$$A_6 K_6 + A_7 K_7 + A_8 K_8,$$

be such a function (A_6, A_7, A_8 being of the second order at most), we must

have

$$A_6 W_2(K_6) + A_7 W_2(K_7) + A_8 W_2(K_8) = 0,$$
$$A_6 W_3(K_6) + A_7 W_3(K_7) + A_8 W_3(K_8) = 0.$$

Hence A_6, A_7, A_8 must be proportional to three quantities of which the first is

$$(15)\quad \frac{1}{\begin{bmatrix}2\,3\,4\,5\,7\\1\,2\,3\,4\,5\end{bmatrix}\begin{bmatrix}2\,3\,4\,5\,8\\1\,2\,3\,4\,5\end{bmatrix}} \cdot \frac{\begin{bmatrix}1\,2\,3\\1\,2\,3\end{bmatrix}^4}{\begin{bmatrix}1\,2\\1\,2\end{bmatrix}^8\begin{bmatrix}1\,3\\1\,2\end{bmatrix}^8} \cdot \begin{vmatrix}\begin{bmatrix}1\,2\,3\,4\,5\,7\\1\,2\,3\,4\,5\,6\end{bmatrix}\begin{bmatrix}1\,2\,3\,4\,5\,8\\1\,2\,3\,4\,5\,6\end{bmatrix} \\ \begin{bmatrix}1\,2\,3\,4\,5\,7\\1\,2\,3\,4\,5\,7\end{bmatrix}\begin{bmatrix}1\,2\,3\,4\,5\,8\\1\,2\,3\,4\,5\,7\end{bmatrix}\end{vmatrix},$$

and the others are obtained from this by a cyclic interchange of the subscripts 6, 7, and 8. Now under the entire group the magnifying factors of the relative invariants

$$\begin{bmatrix}2\,3\,4\,5\,7\\1\,2\,3\,4\,5\end{bmatrix}, \begin{bmatrix}2\,3\,4\,5\,8\\1\,2\,3\,4\,5\end{bmatrix}, \begin{bmatrix}1\,2\,3\\1\,2\,3\end{bmatrix}, \begin{bmatrix}1\,2\\1\,2\end{bmatrix}, \begin{bmatrix}1\,3\\1\,2\end{bmatrix}$$

are respectively

$$\frac{(\Phi_x \Psi_y - \Psi_x \Phi_y)^{10}}{\prod_{\iota=2\,3\,4\,5\,7}(\Phi_x + \Phi_y y_\iota')^4}, \quad \frac{(\Phi_x \Psi_y - \Psi_x \Phi_y)^{10}}{\prod_{\iota=2\,3\,4\,5\,8}(\Phi_x + \Phi_y y_\iota')^4}, \quad \frac{(\Phi_x \Psi_y - \Psi_x \Phi_y)^3}{\prod_{\iota=1\,2\,3}(\Phi_x + \Phi_y y_\iota')^2},$$

$$\frac{\Phi_x \Psi_y - \Psi_x \Phi_y}{\prod_{12}(\Phi_x + \Phi_y y_\iota')}, \quad \frac{\Phi_x \Psi_y - \Psi_x \Phi_y}{\prod_{13}(\Phi_x + \Phi_y y_\iota')}.$$

The four 6-rowed determinants are not relative invariants, but

$$\begin{bmatrix}1\,2\,3\,4\,5\,7\\1\,2\,3\,4\,5\,6\end{bmatrix}, \quad \begin{bmatrix}1\,2\,3\,4\,5\,7\\1\,2\,3\,4\,5\,7\end{bmatrix}$$

become respectively

$$\frac{(\Phi_x \Psi_y - \Psi_x \Phi_y)^{11}}{\prod_{\iota=2\,3\,4\,5\,7}(\Phi_x + \Phi_y y_\iota')^4} \cdot \left\{\Phi_x \begin{bmatrix}1\,2\,3\,4\,5\,7\\1\,2\,3\,4\,5\,6\end{bmatrix} + \Phi_y \begin{bmatrix}1\,2\,3\,4\,5\,7\\1\,2\,3\,4\,5\,7\end{bmatrix}\right\},$$

$$\frac{(\Phi_x \Psi_y - \Psi_x \Phi_y)^{11}}{\prod_{\iota=2\,3\,4\,5\,8}(\Phi_x + \Phi_y y_\iota')^4} \cdot \left\{\Psi_x \begin{bmatrix}1\,2\,3\,4\,5\,7\\1\,2\,3\,4\,5\,6\end{bmatrix} + \Phi_y \begin{bmatrix}1\,2\,3\,4\,5\,7\\1\,2\,3\,4\,5\,7\end{bmatrix}\right\},$$

and the other two similar expressions are obtained by substituting the subscript 8 for the subscript 7.

The determinant of the last four is therefore a relative invariant with the magnifying factor

$$\frac{(\Phi_x \Psi_y - \Psi_x \Phi_y)^{23}}{\left(\prod_{1\,2\,3\,4\,5\,7}\right)\left(\prod_{1\,2\,3\,4\,5\,8}\right)}.$$

The whole expression (15) is found on multiplication to be a relative invariant of the entire group with the magnifying factor $(\Phi_x \Psi_y - \Psi_x \Phi_y)^{-1}$, and this holds likewise for the two similar expressions. If, then, we take A_6, A_7, A_8 *equal*—and not merely proportional—to the expressions of type (15), we see that

$$A_6 K_6 + A_7 K_7 + A_8 K_8 \tag{16}$$

is an absolute invariant of the area-preserving group. It satisfies, moreover, the equations $W_2 f = 0$, $W_3 f = 0$.

There remains only the equation $W_1 f = 0$; this shows that our invariant must be homogeneous of degree zero in all of the variables. The expression (16) is homogeneous of degree -2. J_5 is an absolute invariant of the area-preserving group, and satisfies the equations $W_2 f = 0$, $W_3 f = 0$. It is homogeneous and of degree $-\frac{1}{3}$. The expression

$$J_5^{-6}(A_6 K_6 + A_7 K_7 + A_8 K_8)$$

is therefore homogeneous of degree zero, and is a third-order invariant of the entire group of point transformations.

We do not wish to enter here into the question of what we may regard as limiting forms of the configurations we have just considered, i. e., those configurations some of whose elements have contact of higher than zero order. Rabut has discussed a certain class of them for the group of all point transformations. There are, however, limiting configurations even of the second order possessing invariants under this group which he has failed to detect.

6. **The equilong group.** The imprimitive group

$$X = \Phi(x), \qquad Y = y\Phi'(x) + \chi(x)$$

is known as the equilong group of transformations* when x and y are interpreted as Hessian line coördinates, and has interesting geometric properties.†

We confine ourselves to point coördinates. It is clear from the result of Section 4 that three elements having the same abscissa but different ordinates for the points at which they are situated have an infinite series of invariants.

In the consideration of this group we shall depart from the method of infinitesimal transformations; it is a little more convenient to use the method of elimination.‡ This process consists in successive differentiation of the finite equations of the group, and elimination from the resulting equations of the arbitrary functions. A relation of the nth order of the form

$$f(x, y_i, y_i', \cdots, y_i^{(n)}; X, Y_i, Y_i', \cdots, Y_i^{(n)}) = 0$$

* Scheffers, Mathematische Annalen, vol. 60.

† See papers by Kasner already cited, and also P. H. Linehan, *Contributions to equilong geometry*, Dissertation, Columbia University (1915).

‡ Lie, *Theorie der Transformationsgruppen*, vol. 1, Kap. 13, § 58.

being obtained, the separation of the variables is effected by replacing all the variables but one of the transformed configuration by constants, and then solving for the remaining one. The result is an invariant function of the variables of the first configuration. Care must be taken in the choice of the constants to see that the configuration corresponding to the chosen values has no special properties under the group.

To carry out the operations we have indicated, it is convenient to choose the equations in the form

$$X = \int e^{-\Phi(x)}\,dx, \qquad Y = ye^{-\Phi(x)} + \chi_1(x).$$

Differentiating n times in succession we get

$$\begin{aligned} Y' &= y' - \Phi' y + \chi_2, \\ Y'' &= e^{\Phi}\{y'' - (\Phi' y' + \Phi'' y)\} + \chi_3, \\ &\cdots\cdots\cdots\cdots, \\ Y^{(n)} &= e^{\Phi}\frac{d}{dx}(Y^{(n-1)}). \end{aligned}$$

The substitution of the constants for $Y_\iota, Y_\iota, \cdots, Y_\iota^{(n)}$ may be conveniently performed before the elimination of the arbitrary functions. We choose $Y_1 - Y_2 = 1$, $Y_1^{(k)} - Y_2^{(k)} = 0$ for all values of k from 1 to n. This restriction on two of the elements of the transformed configuration leaves that configuration perfectly general under the group. We have to solve the following equations for the arbitrary functions $e^{\Phi}, \Phi', \Phi'', \cdots, \Phi^{(n)}$.

$$\begin{aligned} 1 &= e^{-\Phi}(y_1 - y_2), \\ 0 &= y_1' - y_2' - \Phi'(y_1 - y_2), \\ 0 &= y_1'' - y_2'' - \Phi'(y_1' - y_2') - \Phi''(y_1 - y_2), \\ &\cdots\cdots\cdots\cdots, \\ 0 &= y_1^{(n)} - y_2^{(n)} - \Phi'(y_1^{(n-1)} - y_2^{(n-1)}) - c_1\Phi''(y_1^{(n-2)} - y_2^{(n-2)}) - \cdots. \end{aligned} \tag{17}$$

A great advantage of this particular choice of the constants now appears. Each of the equations (17) might be obtained by direct differentiation from the preceding, therefore when we have expressed $e^{\Phi}, \Phi', \cdots, \Phi^{(n)}$ in terms of $y_\iota^{(k)} - y_2^{(k)}$ by solving these equations we see that $\Phi'', \Phi''', \cdots, \Phi^{(n)}$ are such as follow from Φ'' by successive differentiation, and that $\Phi' e^{\Phi}$ is simply the derivative of e^{Φ}. We have then

$$e^{\Phi} = y_1 - y_2, \qquad \Phi' = \frac{y_1' - y_2'}{y_1 - y_2}, \qquad \Phi'' = \frac{d\Phi'}{dx}, \qquad \cdots. \tag{18}$$

We have to substitute these values successively in the equations

$$
\begin{aligned}
&Y_1 - Y_3 = e^{-\Phi}(y_1 - y_3),\\
&Y_1' - Y_3' = y_1' - y_3' - \Phi'(y_1 - y_3),\\
(19)\qquad &Y_1'' - Y_3'' = e^{\Phi}\{y_1'' - y_3'' - \Phi'(y_1' - y_3') - \Phi''(y_1 - y_3)\},\\
&\cdots\cdots\cdots\cdots,\\
&Y_1^{(n)} - Y_3^{(n)} = e^{\Phi}\frac{d}{dx}(Y_1^{(n-1)} - Y_3^{(n-1)}).
\end{aligned}
$$

The expressions on the right are the invariants of the different orders. But from the way in which the constants were chosen we see that the operation by which the equations (17) are successively derived might just as well be performed after the substitution of the values (18) as before. In general, then

$$I_k = (y_1 - y_2)\frac{dI_{k-1}}{dx},$$

where I_k is the invariant of order k, while I_{k-1} is that of order $k-1$. The first of the series is

$$I_0 \equiv \frac{y_1 - y_3}{y_1 - y_2}.$$

III. Contact transformations

1. **Finite groups.** Since "contact of zero order" is not preserved by contact transformations in general, no distinction can be made between configurations whose elements have contact of zero order and those whose elements are situated at different neighboring points. With this exception, the results obtained for finite groups of point transformations hold also for finite groups of contact transformations. The fundamental configuration consists of two elements through two distinct points $(x_1\, y_1)$ and $(x_2\, y_2)$; and the number of independent mixed invariants is equal to the number of essential parameters in the group.

2. **Infinite groups. Single element.** In the case of infinite groups single elements may have invariants of order zero or 1. If an invariant of order zero exists, $f(x, y)$ say, we have

$$\Omega(x, y, X, Y) \equiv f(X, Y) - f(x, y) = 0.$$

Then the determinant

$$
(1)\qquad \begin{vmatrix}
0, & \dfrac{\partial\Omega}{\partial x}, & \dfrac{\partial\Omega}{\partial y}\\[2ex]
\dfrac{\partial\Omega}{\partial X}, & \dfrac{\partial^2\Omega}{\partial x\partial X}, & \dfrac{\partial^2\Omega}{\partial y\partial X}\\[2ex]
\dfrac{\partial\Omega}{\partial Y}, & \dfrac{\partial^2\Omega}{\partial x\partial Y}, & \dfrac{\partial^2\Omega}{\partial y\partial Y}
\end{vmatrix}
$$

vanishes identically. We have then the

THEOREM. *A contact transformation having an invariant of order zero is merely an extended point transformation.*

All groups of contact transformations leaving invariant a differential equation of the second order are reducible to the group of point transformations. Hence the

THEOREM. *A contact transformation having an invariant of the first order is reducible to the form*

$$X = x, \qquad Y = \Psi(x, y).$$

Finally, we know that no infinite group of contact transformations can leave invariant a differential equation of the third or higher order, and we have the

THEOREM. *No infinite group of contact transformations can have an invariant of a single element of the second or higher order.*

3. **Infinite groups. Rabut's Theorem.** Every infinite group of contact transformations has, however, invariants of certain configurations. The group of all contact transformations was treated by Rabut, who showed that it had no essential invariants other than those of individual bunches, and, moreover, that all the elements must have contact of at least the first order. Under these circumstances he proved the

THEOREM. *The invariants of the group of all contact transformations are simply those of the group of all point transformations, except that the order of each derivative is increased by unity.*

We shall verify this theorem by comparing the equivalent complete systems of the two groups.

The general infinitesimal transformation of the group of all contact transformations is

$$\delta x = \xi(x, y, y')\,\delta t, \qquad \delta y = \eta(x, y, y')\,\delta t, \qquad \delta y' = \pi(x, y, y')\,\delta t,$$

where ξ, η, π are independent functions of x, y, y' expressible, however, in terms of the derivatives of one arbitrary function $W(x, y, y')$. We have

$$\xi = W_{y'}, \qquad \eta = y' W_{y'} - W, \qquad \pi = -W_x - W_y y'.$$

The increment of y'' is

$$\eta'' = (\pi_x + \pi_y y') + y''\{\pi_{y'} - (\xi_x + \xi_y y')\} - y''^2 \xi_{y'}.$$

Compare this with the increment of y' under the group of all point transformations. The rule of operation for the successive increments is always for $n > 1$

$$\eta^{(n)} = \frac{d\eta^{(n-1)}}{dx} - y^{(n)} \frac{d\xi}{dx},$$

exactly as in the case of the point transformations. The only difference is that instead of the arbitrary functions of two variables $f(x, y)$, $f_x(x, y)$, $f_y(x, y)$, we now have $F(x, y, y')$, $F_x(x, y, y') + y' F_y(x, y, y')$, and $F_{y'}(x, y, y')$, and that $y^{(k+1)}$ always takes the place of $y^{(k)}$. We thus get the same complete system for order $n + 1$ as in the case of the point transformations for order n, except that $y^{(k+1)}$ takes the place of $y^{(k)}$. We find that $2n + 2$ curves have an invariant of order $n + 1$, obtained from the corresponding invariant of order n for the point transformations by the above substitution.

4. **Certain reducible groups.** We noticed an exception to the general rule that each bunch of elements had its own invariants in the case of infinite groups of point transformations. A similar exception occurs in the case of a certain class of contact transformations, viz., those which leave invariant a differential equation of the second order in a special way: if $u(x, y, y')$ be a first integral, they transform the ∞^1 families $u = c$ among themselves. If we choose such a configuration that the first three coördinates x, y, y' of each element satisfy a relation $u(x, y, y') = c$, that is, the elements have first-order contact with different curves of the *same family* of integral-curves, the function $u(x, y, y')$ will be the same for each bunch. An arbitrary function $\lambda(u)$ occurring in the extended infinitesimal transformations of the group will lead to equations in the equivalent complete system among variables from all the bunches. It is sometimes possible under these circumstances for the configuration to have invariants other than those of individual bunches.

It should be noticed that if the function u does not involve y', the group which leaves the totality of the ∞^1 families $u = c$ invariant is merely a group of point transformations. For we have a relation of the type

$$\Omega(x, y, X, Y) \equiv u(X, Y) - f\{u(x, y)\} = 0,$$

which is such that the determinant (1) vanishes.

Now the group of contact transformations which leaves invariant a given differential equation of the second order is reducible to the group of point transformations. All such groups are of the form

$$u(X, Y, Y') = \Phi\{u(x, y, y'), v(x, y, y')\},$$
$$v(X, Y, Y') = \Psi\{u(x, y, y'), v(x, y, y')\},$$

where u and v are independent first integrals of the given differential equation. If the group has the special property of leaving the totality of the ∞^1 families $u = c$ invariant, it has the form

$$(2) \qquad \begin{aligned} u(X, Y, Y') &= \Phi\{u(x, y, y')\}, \\ v(X, Y, Y') &= \Psi\{u(x, y, y'), v(x, y, y')\}. \end{aligned}$$

If we transform this group by the contact transformation

$$x = u(x, y, y'),$$

$$y = v(x, y, y'),$$

we obtain the group of point transformations

$$X = \Phi(x),$$

$$Y = \Psi(x, y).$$

It is now clear that the only groups possessing essential invariants of more than one bunch of elements are those for which $\Psi(u, v)$ has the special form

$$\frac{\Psi_1(u) + v\Psi_2(u)}{\Psi_3(u) + v\Psi_4(u)}.$$

These are all reducible by the transformation (2) to the group of point transformations of Section 3, Part II. They have an invariant of the first order of four elements

$$\frac{(v_1 - v_2)(v_3 - v_4)}{(v_2 - v_3)(v_4 - v_1)},$$

and thereafter a series of invariants, one of each order, for five curves.

A simple illustration is the largest subgroup of the group preserving parallelism. Here $u = y'$, $v = y - xy'$, and the group has the form

$$Y' = \Phi(y'),$$

$$Y - XY' = \frac{\Psi_1(y') + \Psi_2(y') \cdot (y - xy')}{\Psi_3(y') + \Psi_4(y') \cdot (y - xy')}.$$

This is reduced to a group of point transformations by the contact transformation

$$X = y',$$

$$Y = y - xy',$$

$$Y' = -x.$$

The group has the invariant of four parallel elements

$$\frac{\{y_1 - y_2 + (x_1 - x_2)y'\}\{y_3 - y_4 + (x_3 - x_4)y'\}}{\{y_2 - y_3 + (x_2 - x_3)y'\}\{y_4 - y_1 + (x_4 - x_1)y'\}}.$$

Five elements whose tangents are parallel have a series of invariants one of each order $n > 1$.

5. **Number of elements having invariants.** Finally, the argument of Section 4, Part II, is applicable to contact transformations, and we obtain the

THEOREM. *If the group involves arbitrary functions of more than one variable, λ_n is an increasing function of n; otherwise it attains a final value $k+1$ (where k is the number of arbitrary functions of one variable in the unextended form of the group), after a certain finite value of n.*